

**Multikonferenz  
Wirtschaftsinformatik (MKWI)  
2016**

**Research-in-Progress- und Poster-Beiträge**

Technische Universität Ilmenau  
09.-11. März 2016



# Systematischer Einsatz von Playful Prototyping und Gamification im Requirements Engineering

Rainer Alt<sup>1</sup>, Kai-Christoph Hamborg<sup>2</sup>, Franziska Metzger<sup>3</sup>, Tammo Straatmann<sup>2</sup> und Alexander Hofmann<sup>3</sup>

<sup>1</sup> Universität Leipzig, Institut für Wirtschaftsinformatik, rainer.alt@uni-leipzig.de

<sup>2</sup> Universität Osnabrück, Institut für Psychologie, {khamborg | tstraatm}@uni-osnabrueck.de

<sup>3</sup> MaibornWolff GmbH, München, {franziska.metzger | alexander.hofmann}@maibornwolff.de

## Abstract

Das Requirements Engineering erlebt durch Entwicklungen wie die Konsumentenorientierung oder die Vernetzung von Organisationen ein verändertes Umfeld, das eine breitere und durchgängigere Einbindung von Beteiligten in der Anforderungsermittlung erfordert. Dieser Beitrag stellt dazu einen Ansatz vor, der spielerische Elemente und Elemente der Gamification für die Anforderungsanalyse integriert. Das im Workshopformat eingesetzte Playful Prototyping zielt, vermittelt über positiven Affekt, auf die Förderung von Kreativität und Motivation beteiligter Personen als Voraussetzung für die Entwicklung innovativer Anforderungen. Dieses Vorgehen flankiert eine „Gamified“ RE-Umgebung (GameReq), welche die Anforderungen aus dem Playful Prototyping verwaltet und die Analyse mit einer Crowd-Plattform um weitere Stakeholdergruppen erweitert.

## 1 Requirements Engineering im Wandel

In den vergangenen Jahren hat die Anforderungsermittlung bzw. das Requirements Engineering (RE) zahlreiche grundlegende Veränderungen erfahren. Zunächst besitzen Wirtschaftsgüter in zunehmendem Maße eine Softwarekomponente, sodass die Nutzung von Software eine gesamtwirtschaftliche Bedeutung besitzt. Je adäquater die Software die Aufgaben aus Nutzer- und Ergebnissicht abbildet, desto stärker ist von einem positiven Beitrag zur „Usability“ und einer effizienten Aufgabendurchführung auszugehen. Dabei beziehen die Entwicklungen nicht mehr nur die internen Mitarbeiter eines Unternehmens, sondern zunehmend auch externe Partner und Endkunden („Consumerization“, Weiß & Leimeister, 2012) mit ein. Mit der Vernetzung von Organisationen, Individuen und Dingen („Hyperconnectivity“, Heisterberg & Verma, 2014) entstehen komplexe Wirkungsgeflechte, die in der Anforderungsanalyse zu erfassen und verstehen sind. Eine besondere Herausforderung besteht für den RE-Prozess zudem darin, dass bei der Anforderungserfassung sowohl die Erhebung bestehender als auch neuer Abläufe und innovativer Funktionalitäten gefordert ist.

Klassische Verfahren im Anforderungsermittlungsprozess (RE-Prozess) haben zwar in den vergangenen Jahren eine Weiterentwicklung hin zu agilen und nutzerorientierten Vorgehensweisen erfahren, die starre, dokumentationsgetriebenen Modelle mit sequentielltem Ablauf ablösen. Ebenso zeigen jüngere Ansätze des „Crowdsourcing“ (Hammon & Hippner, 2012) insbesondere im Bereich der Entwicklung von Open Source Software einen Ansatz zur Partizipation interner und externer Anspruchsgruppen („Stakeholder“), der häufig auch motivierende Elemente der Gamification vorsieht. Beispielsweise geben Leaderboards und Badges Hinweise auf die leistungsfähigen Akteure einer Community.

Allerdings ist die beteiligungsorientierte Erhebung kreativer Anforderungen für die Entwicklung innovativer Anwendungen bislang weder mit klassischen noch agilen Softwareentwicklungsmethoden verbunden. Umgekehrt fehlt dem Crowdsourcing die Anbindung mit RE-Methoden. Vor diesem Hintergrund beschreibt der vorliegende „Research-in-Progress“-Beitrag einen in Entwicklung befindlichen innovativen Ansatz, welcher mittels aufeinander abgestimmter physischer und virtueller Ebenen spielerische Elemente in den RE-Prozess einbettet und damit Nutzer zur kreativen Mitarbeit motiviert.

## **2 Grenzen und Herausforderungen bestehender RE-Ansätze**

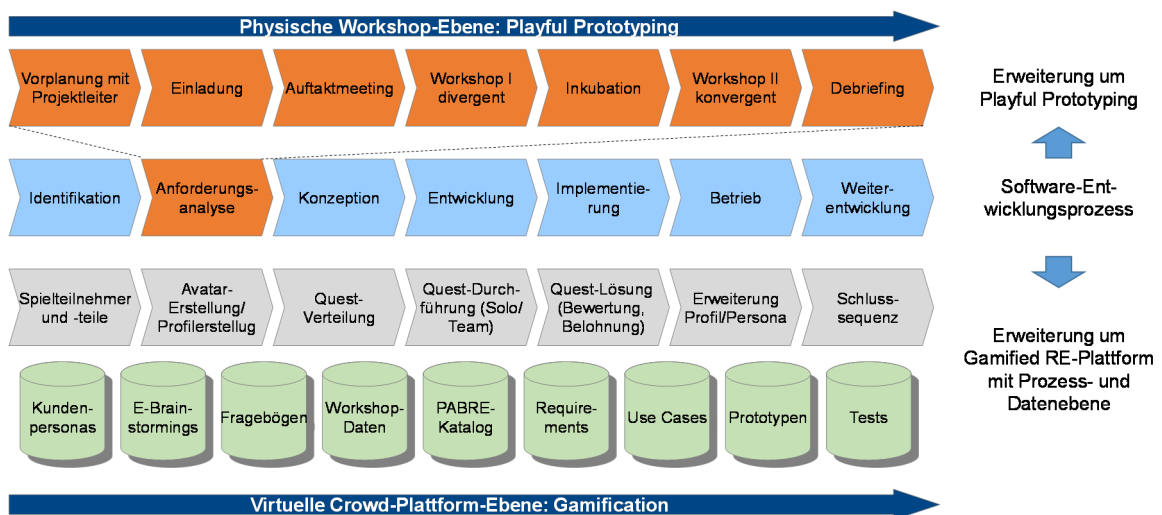
Zur Software-Entwicklung haben sich in Literatur und Praxis Modelle für die Strukturierung des Entwicklungsprozesses etabliert (z.B. Schatten et al., 2010). Der Anforderungsermittlung kommt in diesen Modellen ein zentraler Stellenwert zu, da hier die Definition von Vorgaben für sämtliche nachfolgende Phasen erfolgt (Balzert, 2009). Unterlassungen in dieser Phase können zu aufwändigen Nachbesserungen in späteren Projektphasen, zu hohen Folgekosten oder gar zum Scheitern von Projekten führen (Wallmüller, 2011). Als Hauptursachen für schlecht verlaufene oder gescheiterte Softwareentwicklungsprojekte gelten unvollständige und inkonsistente Anforderungen, sowie Kommunikationsprobleme mit Kunden und innerhalb von Entwicklerteams (Goll, 2011). Der systematischen und effektiven Einbeziehung externer Anspruchsgruppen (Kunden, Nutzer etc.) kommt daher besondere Bedeutung für die Erschließung technologischer Potenziale (Mobile Geräte, Cloud Computing, Social Media, Big Data etc.), die Entwicklung innovativer Lösungen und für die Akzeptanz der späteren Softwarelösung zu.

Im klassischen RE-Prozess stehen dem jedoch wichtige Barrieren gegenüber: einerseits bezieht die Anforderungserfassung für betriebliche Informationssysteme typischerweise nur wenige Nutzer ein, andererseits sind viele Teilnehmer für die strukturierte Anforderungserhebung oft weder motiviert noch kreativ bei der Entwicklung von Gestaltungsideen. Als drittes Defizit sind zwischen den Ergebnissen aus Anforderungserfassungs-Workshops und dem sich anschließenden Entwicklungsprozess häufig Medienbrüche zu beobachten, die auf einen Mangel an einer durchgängigen Dokumentation der Anforderungen hinweisen. Zudem entsprechen bisherige Vorgehensweisen der Anforderungsermittlung nicht dem Prinzip agiler Softwareentwicklung, nach dem die Erhebung von Anforderungen ein begleitender Prozess in der Umsetzung ist.

Besonders für die Entwicklung innovativer Anwendungen besteht eine Herausforderung darin, dass sich Anforderungen nicht alleine aus der Analyse und Deskription des Nutzungskontexts ableiten lassen, sondern einen iterativen, konstruktiv-kreativen Prozess erfordern (Rogers et al., 2011), der ausgehend von einem gegebenen Ist-Zustand einen nicht klar definierten Soll-Zustand, das neue IT-Produkt, anstrebt. Die aufgezeigten Problempunkte und Herausforderungen soll der nachfolgend dargestellte RE-Ansatz adressieren.

### 3 Forschungsansatz

Die vorliegende Forschung zielt auf eine Verzahnung von zwei Elementen in der Anforderungsanalyse: Während das Playful Prototyping die kreative Entwicklung von Anforderungen durch den Einsatz spielerischer Elemente ermöglicht, ergänzt Gamification eine entlang des gesamten RE-Prozess genutzte Crowd-Plattform. Die Gesamtmethodik soll auf möglichst beliebige Gestaltungsprojekte anwendbar sein und insbesondere zur Entwicklung innovativer Softwareprodukte beitragen. Die erste Konzeption einer Lösungsarchitektur orientiert sich am klassischen RE-Prozess, der von der Bedarfsidentifikation für eine neue Software über die Anforderungsanalyse, die Konzeption, die Entwicklung, die Implementierung und den Betrieb hin zur Weiterentwicklung erstreckt (s. Bild 1). Das Konzept sieht eine physische Ebene für die intensive Anforderungsermittlung in einem kleinen Kreis beteiligter Personen in Workshops vor. Eine virtuelle Ebene übernimmt die Vor- und Nachbereitung der physischen Ebene sowie die Erweiterung um weitere Nutzerkreise im Sinne bekannter Crowdsourcing-Ansätze (Leimeister et al., 2009).



**Bild 1: RE-Erweiterung mit Playful Prototyping und Gamification**

#### 3.1 Konzept des Playful Prototyping

Wird die Entwicklung von Anwendungen als konstruktiv-kreativer Problemlösungsprozess verstanden, besteht eine Herausforderung darin, Konzepte zu finden, wie sich die hierfür erforderliche Kreativität und auch Motivation der beteiligten Stakeholder im RE-Prozess fördern lässt. Einen Ansatzpunkt bieten Befunde, nach denen positiver Affekt die Kreativität in Problemlöseprozessen fördert (De Dreu et al., 2008). Das Konzept des Playful Prototyping nimmt diese Befunde auf. Integraler Bestandteil dieses Ansatzes ist der Einsatz spielerischer Elemente, um positiven Affekt (Mainemelis & Ronson, 2006) und damit Kreativität auf Seiten der im RE-Prozess beteiligten Personen zu evozieren. Zudem wird davon ausgegangen, dass Spielsituationen das Interesse der Beteiligten wecken (Kirriemuir & McFarlane, 2004) und es erleichtern, sich gedanklich von den Grenzen des Alltags, d.h. von funktionalen Fixierungen, zu lösen und unterschiedliche Perspektiven einzunehmen, worin wiederum eine Voraussetzung für kreatives Problemlösen besteht (Grant & Berry, 2011). Schließlich geht das Vorgehen davon aus, dass die Motivation der Beteiligten im RE-Prozess durch die Möglichkeit angeregt wird, mit einem

gewissen Ausmaß an Selbstbestimmtheit in Spielsituationen physisch aktiv sein zu können (Ryan et al., 2009).

Ausgangspunkt für die Methodik des Playful Prototyping ist die Formulierung einer allgemeinen Gestaltungsvision für ein zu entwickelndes System und die Analyse des Nutzungskontexts mit Identifikation von Nutzergruppen, Kernaufgaben und Handlungszielen. Für das weitere Vorgehen werden zwei allgemeine Phasen unterschieden (Johannesson & Perjons, 2014):

1. Eine kreative, **divergente** Phase zur Generierung möglichst vieler valider Anforderungen. Einen Rahmen bieten Nutzungsszenarien bzw. Prototypen zur Identifikation von Anforderungen (Go & Carroll, 2004), die Teilnehmer in einem Workshop-Format mittels Personas (Cooper et al., 2014), durchspielen. Kernelemente der Spielsituation sind:

- der *Playground* als physische Plattform für das Durchspielen von Nutzungsszenarien,
- *Personas* als abstrahierte aber gegenständlich dargestellte Nutzer, mit denen Nutzungsszenarien auf dem Playground durchgespielt werden,
- die *User Journey*, die das Durchspielen von Nutzungsszenarien mit Personas bezeichnet. Eine User Journey ist durch einen, für den Nutzungskontext typischen, aufgabenbezogenen Start- und Zielpunkt gekennzeichnet. Im Rahmen kreativ gestaltbarer User Journey entwickeln die Teilnehmer Ideen, wie eine Anwendung das aufgabenbezogene Handeln unterstützen kann,
- *Touchpoints*, welche die für eine effektive, effiziente und zufriedenstellende User Journey notwendigen Interaktionen mit der Anwendung kennzeichnen,
- *User Stories* und *Story-Cards* (Wirdemann, 2011) für die Dokumentation der resultierenden Anforderungen an die Anwendung.

2. Eine sich anschließende **konvergente** Phase dient der Ideenkonsolidierung der in Form von User-Stories dokumentierten Anforderungen mit einer an die Claims Analysis angelehnten Bewertungsmethodik unter Berücksichtigung definierter Akzeptanzkriterien.

### 3.2 Crowd-Plattform GameReq

Der physische Teil der Anforderungsermittlung ist mit einer Crowd-Plattform („Gamified RE-Plattform“, GameReq) abgestimmt, der bereits im Vorfeld des Workshops die Teilnehmer in den spielerischen Kontext einbettet und durch eine virtuelle Repräsentation der Personas deren Weiterleben über den Zeitrahmen des Zusammentreffens in den Workshops hinaus ermöglicht. Intrinsische Motivation, die verbunden mit persönlichen Zielen und Überzeugungen nachhaltiger ist, kann durch eine kohärente Player Journey und die nichtlineare und weit angelegte Entwicklung des vom Spieler gespielten Helden (Avatar) verstärkt werden (Lombriser et al., 2011). Für eine GameReq-Plattform bestehen verschiedene Vorarbeiten:

- Die Gamification-Literatur schlägt potenzielle Elemente vor. Dazu zählen (Deterding et al., 2011): *Avatare* (virtuelle Spielfiguren entlang der Personas), *Quests* (zu erledigende (Teil)Aufgaben im Softwareentwicklungsprozess), *Punkte* (Belohnung für die Lösung von Quests), *Badges* (Auszeichnungen für besondere Aktionen des Avatars) sowie *Items* (Gegenstände des Avatars, die z.B. seinen Status erhöhen).
- Ansätze des RE zur Dokumentation von Anforderungen in standardisierter Form für nachfolgende Phasen des Softwareentwicklungsprozesses. Beispiele sind der Katalog von

Requirement-Patterns PABRE (Renault et al., 2009) und Standards wie das Requirements Interchange Format (ReqIF) der Object Management Group (OMG).

- Crowd-Plattformen zum Einsatz im RE (Renzel et al., 2013; Snijders et al., 2015).

Die Verbindung von physischer und virtueller Ebene sollen typische Probleme von Anforderungs-Workshops, wie begrenzte Gruppengrößen, überwinden. Zwar haben Gamification-Ansätze bereits auf die Potenziale für den RE-Prozess hingewiesen, jedoch sind bislang primär einfache Spielelemente zum Einsatz gekommen (Snijders et al., 2015; Rapp, 2013) und bei der Entwicklung neuer Ideen und Anforderungen bestehen noch Defizite (Marshburn & Henry, 2013). Ferner sind gegenüber extrinsischen Motivatoren (z.B. Punktevergaben) implizite Motivatoren des Gamedesign, die tiefer in der Spielmechanik liegen, wie User-Stories und Personas für den Einsatz in Gamification noch wenig anzutreffen.

### 3.3 Erste Erfahrungen

Eine erste Version der Playful Prototyping-Methodik wurde mit Fach- und Methodenexperten eines Software-Entwicklungsunternehmens erprobt. Auf physischer Ebene kamen Personas zum Einsatz, welche die Teilnehmer mit Lego und weiteren Materialien modelliert und Handlungsszenarien auf Playgrounds erstellt haben. Die Kontaktpunkte der Personas zum Handlungsszenario wurden markiert und Anforderungen in Form von User-Stories auf Story-Cards festgehalten. Jedem Schritt folgte eine Evaluation des Vorgehens in gemeinsamen Reflexionsphasen und parallel dazu eine Erhebung der Anforderungen an die GameReq-Plattform. Für Begeisterung hat gesorgt, dass die Haptik und das Verwenden der Materialien den kreativen Prozess verstärkt haben, so dass Ideen entstanden, die nach eigener Aussage der Gruppe ohne diese nicht aufgekommen wären. Insgesamt hat das von den bisherigen Anforderungs-Workshops abweichende Format für ein höheres Engagement der Teilnehmer und nach Einschätzung der Teilnehmer für bessere Lösungen gesorgt. Alle Teilnehmer begrüßten zudem die geplante Fortführung der Anforderungserhebung in einem Tool, das die entwickelten Personas und Handlungsszenarien aufgreift.

## 4 Zusammenfassung und Ausblick

Die beschriebene „Research-in-Progress“ verbindet physische und virtuelle Interaktionselemente mit einer Crowd-Plattform, um den RE-Prozess kreativer und durchgängiger zu gestalten und breitere Nutzergruppen einzubeziehen. Der interdisziplinäre Ansatz nutzt Kenntnisse aus der Psychologie, um Bedingungen für Kreativität und Motivation im RE-Prozess zu schaffen. Seitens der Wirtschaftsinformatik fließen Forschungsergebnisse aus dem Bereich der Gamification ein, welche mittels einer Crowd-Plattform die Motivation vieler Beteiligter entlang des gesamten RE-Prozesses beibehält und gleichzeitig auf eine systematische Dokumentation abzielt.

Der Forschungsansatz setzt am Defizit an, dass bestehende RE-Werkzeuge die frühe Phase der Anforderungsermittlung unter den künftigen Rahmenbedingungen nicht ausreichend unterstützen. So schaffen agile Vorgehensweisen zwar Vorteile bezüglich Geschwindigkeit und Interaktion, jedoch berücksichtigen sie nur wenig den innovations- und motivationsfördernden Charakter, die Integration weiterer Nutzerkreise und die systematische Dokumentation für den Softwareentwicklungsprozess. Zumeist werden Anforderungen in der Erhebung erst in natürlicher Sprache mit dem Vorteil leichter Verständlichkeit für alle Beteiligten formuliert und im späteren Projektverlauf in formale Darstellungen, wie UML-Diagramme transformiert. Die Resultate sind händisch einzupflegen, was einen zeitintensiven und fehleranfälligen Arbeitsschritt darstellt.

Weiterhin beziehen sich bestehende marktreife Produkte ausschließlich auf das Management bereits erhobener Anforderungen, unterstützen aber nicht den Erhebungsprozess.

Mittels des skizzierten dualen Einsatzes spielerischer Elemente lässt sich die Erhebung von Anforderungen in der frühen Phase der Anforderungsermittlung auf einer physischen und auf einer virtuellen Ebene verbinden. So finden Personas sowohl im Kontext des Playful Prototyping als auch im Rahmen einer Crowd-Plattform in Form von Avataren Anwendung. Dies überwindet das in der Literatur anzutreffende komplementäre Verständnis (Deterding et al., 2011) von „Playful Interaction“ und „Gamification“: Während die physische Ebene „Playful Elements“ einsetzt, sind dies in darauf abgestimmter Weise „Gamified bzw. Gameful Elements“ auf virtueller Ebene. Bislang fehlt die wissenschaftliche Fundierung der Wirkweise beider Elemente. Forschungsergebnisse zur Nachhaltigkeit spielerischer Elemente sind erst ansatzweise verfügbar. So führen die auch als „Pointification“ bezeichneten Elemente der Gamification, wie Punkte, Auszeichnungen und Bestenlisten flüchtige extrinsische Motivation, wodurch sie für komplexe und langfristige Projekte eher ungeeignet erscheinen. Die Verwendung von GameReq soll insbesondere bei größeren Nutzergruppen frühe Akzeptanztests und ein automatisiertes Testen in späteren Entwicklungsphasen unterstützen. Zudem kann die vorgesehene Katalogisierung und ermöglichte dadurch ermöglichte Wiederverwendung und Anpassung von User-Stories zu Effizienzvorteilen in späteren Phasen des Softwareentwicklungsprozesses (z.B. den Tests) führen.

Für den vorgestellten Ansatz soll nach erfolgreicher Verprobung in der Praxis eine psychologisch fundierte RE-Methodik mit einem prototypisch entwickelten Tool, das den Prozess begleitend einrahmt, und ein Werkzeugkasten („Toolbox“) entstehen, der den RE-Prozess effektiv gestaltet.

## 5 Literatur

- Adeptu A, Ahmed KA, Abd YA, Zaabi AA, Svetinovic D (2012). CrowdREquire: A Requirements Engineering Crowdsourcing Platform, in: 2012 AAAI Spring Symposium Series.
- Balzert H (2009). Lehrbuch der Software-Technik - Basiskonzepte und Requirements Engineering, 3. Aufl.. Heidelberg: Spektrum.
- Cohn M (2004). User Stories Applied for Agile Software Development. Boston (MA): Pearson.
- Cooper A, Reimann R, Cronin D, Noessel C (2014). About Face: The Essentials of Interaction Design. Indianapolis (IN): Wiley.
- De Dreu CKW, Baas M, Nijstad BA (2008). Hedonic Tone and Activation Level in the Mood-creativity Link: Toward a Dual Pathway to Creativity Model, in: Journal of Personality and Social Psychology, 94(5), S. 739-756.
- Deterding S, Dixon D, Khaled R, Nacke L (2011). From Game Design Elements to Gamefulness: Defining Gamification. In: Proceedings 15th MindTrek Conference, ACM, S. 9-15.
- Fecher M (2012). Gamification in der Softwareentwicklung: Chancen und Möglichkeiten, Hochschule für Angewandte Wissenschaften Würzburg-Schweinfurt.
- Go K, Carroll JM (2004). The Blind Men and the Elephant: Views of Scenario-based System Design, in: Interactions, 11(6), S. 44-53.
- Goll J (2011). Methoden und Architekturen der Softwaretechnik. Wiesbaden: ViewegTeubner.



- Grant AM, Berry JW (2011). The Necessity of Others is The Mother of Invention: Intrinsic and Prosocial Motivations, Perspective Taking, and Creativity, in: *Academy of Management Journal*, 54(1), S. 73-96.
- Hammon L, Hippner H, (2012). Crowdsourcing, in: *Business & Information Systems Engineering*, 4(3), S. 163-166.
- Heisterberg R, Verma A (Eds) (2014). *Creating Business Agility: How Convergence of Cloud, Social, Mobile, Video, and Big Data Enables Competitive Advantage*. Hoboken (NJ): Wiley.
- Johannesson P, Perjons E (2014). *An Introduction to Design Science*. Heidelberg: Springer.
- Kirriemuir J, McFarlane A (2004). Review in *Games and Learning*, NESTA Futurelab Research Report 8, Bristol: Futurelab.
- Leimeister JM, Huber M, Bretschneider U, Krcmar H (2009). Leveraging Crowdsourcing: Activation-supporting Components for IT-based Ideas Competition, in: *Journal of Management Information Systems*, 26(1), S. 197-224.
- Lombriser P, van der Valk R (2011). *Improving the Quality of the Software Development Lifecycle with Gamification*, Springer: Berlin Heidelberg.
- Mainemelis C, Ronson S (2006). Ideas are Born in Fields of Play: Towards a Theory of Play and Creativity in Organizational Settings, in: *Research in Organizational Behavior*, 27, S. 81-131.
- Marshburn DG, Henry RM (2013). Improving Knowledge Coordination in Early Stages of Software Development Using Gamification, in: *Proceedings Southern Association for Information Systems Conference*, Savannah (GA), S. 123-128.
- Rapp A (2013). Beyond Gamification: Enhancing user Engagement through Meaningful Game Elements. In: *FDG*, S. 485-487, <http://reqif.de/index.php/home.html>, abgerufen am: 4.11.2014.
- Renault S, Mendez-Bonilla O, Franch X, Quer C (2009). PABRE: Pattern-based Requirements Elicitation, in: *Research Challenges in Information Science (RCIS 2009)*, S. 81-92.
- Renzel D, Behrendt M, Klamma R, Jarke M (2013). Requirements Bazaar: Social Requirements Engineering for Community-driven Innovation, in: *RE Conference*, S. 326-327.
- Rogers Y, Sharp H, Preece J (2011). *Interaction Design: Beyond Human-Computer Interaction*, 3. Aufl., Chichester: Wiley.
- Ryan RM, Williams GC, Patrick H, Deci EL (2009). Self-Determination Theory and Physical Activity: The Dynamics of Motivation in Development and Wellness, in: *Hellenic Journal on Psychology*, 6, S. 107-124.
- Schatten A, Biffel S, Demolsky M, Gostischa-Franta E, Östreicher T, Winkler D (2010). *Best Practice Software-Engineering*. Heidelberg: Spektrum.
- Snijders R, Dalpiaz F, Brinkkemper S, Hosseini M, Özum A (2015). Refine: A Gamified Platform for Participatory Requirements Engineering, in: *Proceedings 1st International Workshop on Crowd-Based Requirements Engineering (CrowdRE 2015)*, abgerufen am 4.12.2015.
- Wallmüller E (2011). *Software Quality Engineering*, 3. Aufl., München: Hanser.
- Wirdemann R (2011). *Scrum mit User Stories*, 2. Aufl., München: Hanser.
- Weiß F, Leimeister JM (2012). Consumerization, in: *Wirtschaftsinformatik*, 54(6), S. 351-354.



# Real-Time Business Intelligence: Eine systematische Identifikation aktueller Forschungsbedarfe

Fabian Blaser<sup>1</sup>, Mario Nadj<sup>2</sup> und Christian Schieder<sup>1</sup>

<sup>1</sup> Technische Universität Chemnitz, Professur Wirtschaftsinformatik II, fabian.blaser@s2010.tu-chemnitz.de, christian.schieder@wirtschaft.tu-chemnitz.de

<sup>2</sup> Universität Mannheim, Institut für Enterprise Systems (InES), nadj@es.uni-mannheim.de, SAP SE, Walldorf, mario.nadj@sap.com

## Abstract

Im Zeitalter von Big Data stehen Analytikern und Entscheidungsträgern eine Vielzahl an neuen Datenquellen zur Integration und Erkenntnisgewinnung zur Verfügung. Durch den technologischen Fortschritt der letzten Jahre hat sich der Einsatz der traditionell strategisch fokussierten Business Intelligence (BI) auf die operationale Ebene ausgedehnt. Weil hier häufig zeitkritische Prozesse Ziel der Unterstützung sind, wurde der Begriff der Real-Time Business Intelligence (RTBI) geprägt. Die vorliegende Arbeit widmet sich der systematischen Sammlung, Aufbereitung und Analyse des in letzter Zeit stark angewachsenen Literaturbestands zum Thema RTBI. Ziel dieser Untersuchung ist es, dominierende Forschungsfelder im Bereich RTBI mitsamt ihrer zugrunde liegenden Herausforderungen zu identifizieren und eine Forschungsagenda vorzustellen, welche offene thematische Schwerpunkte und mögliche Ansätze für zukünftige Forschungen in diesen Bereichen präsentiert.

## 1 Einleitung

Zentraler Bestandteil des Konkurrenzkampfes auf marktwirtschaftlich organisierten Märkte ist der Faktor Zeit – sei es bei der Entwicklung neuer Produkte, deren Produktion oder im Vertrieb. Das Unternehmen mit der kürzesten Reaktionszeit kann auf Veränderungen in der Unternehmensumwelt am schnellsten reagieren und somit einen erheblichen Vorteil gegenüber der Konkurrenz realisieren (Stalk 1988). Die Entscheidungsfindung bildet hier häufig einen zeitlichen Flaschenhals. Um diesen Engpass zu überwinden und Entscheidungsträger zu unterstützen, hat sich der Einsatz von Business Intelligence (BI)-Lösungen etabliert. Die Entwicklungen der letzten Jahre, insbesondere im Bereich von Radio Frequency Identification Devices (RFID), von Sensoren sowie der zunehmenden Bedeutung von Social Media stellen Möglichkeiten für die Analyse und Erkenntnisgewinnung mit niedriger zeitlicher Latenz bereit (Geerdink 2013). Ermöglicht durch technologische Fortschritte, nicht zuletzt im Bereich des Complex Event Processings und des Real-Time Data-Warehousing, hat sich die Real-Time Business Intelligence (RTBI) als neuer Ansatz zur Realisierung einer höheren Datenaktualität etabliert. RTBI-Lösungen zeichnen sich durch die Verarbeitung von Daten in Echtzeit und der automatischen Ableitung von Handlungsalternativen

aus. In der Praxis existieren verschiedene Ansätze für RTBI-Systeme, abhängig von der Aktualität der Daten für den jeweiligen Zweck.

In den letzten Jahren ist eine kontinuierliche Zunahme der Forschung im Bereich RTBI zu verzeichnen. Allerdings existiert momentan – nach unserem bestem Wissen – keine vollständige und systematische Forschungsagenda für RTBI. Zweck dieser Arbeit ist somit die Aufarbeitung und Behebung dieses Defizits um dadurch weitere systematische Forschungsansätze in dieser Disziplin zu ermöglichen. Hierbei werden zwei Ziele unterschieden: (1) die Identifikation des State-of-the-Art von RTBI und die Analyse der bestehenden Herausforderungen sowie (2) das Erstellen einer Forschungsagenda. Diese Ziele werden durch die Formulierung der folgenden zwei Forschungsfragen erreicht: (1) *Welches sind die vorherrschenden Forschungsschwerpunkte im Bereich der RTBI?* (2) *Wo liegen potentielle Schwerpunkte für zukünftige Forschungen?*

Nach dieser Einleitung schließt sich in Abschnitt 2 der aktuelle Forschungsstand zum Thema RTBI an. Abschnitt 3 beinhaltet die Erläuterung der angewandten Forschungsmethodik, während die Präsentation der Ergebnisse in Abschnitt 4 erfolgt. In Abschnitt 5 werden mögliche Ansatzpunkte für zukünftige Forschungen beschrieben. Die Konklusion und der Ausblick in Abschnitt 6 bilden den Abschluss dieser Arbeit.

## 2 Real-Time Business Intelligence: Grundlagen und Terminologie

Das weite Begriffsverständnis von BI umfasst alle Methoden und Technologien, die der Speicherung und Bereitstellung operativer Daten sowie deren Analyse und geeigneten Präsentation dienen (Gluchowski et al. 2008). RTBI-Lösungen weisen Architekturen auf, die mit denen traditioneller BI-Systeme eng verwandt sind, aber zusätzliche Komponenten enthalten. Dazu zählen neben einer kontinuierlichen Datenintegration in Echtzeit (Chaudhuri et al. 2011) eine aktive Decision-Engine und eine hochverfügbare Analyseumgebung (McKenna 2011). Eine weitreichende Verknüpfung und Zusammenarbeit zwischen allen RTBI-Komponenten und den zugrundeliegenden Geschäftsprozessen sowie den IT-Systemen des Unternehmens sind notwendig um den erfolgreichen Betrieb einer solchen Lösung zu gewährleisten. Diese Integration wird durch zahlreiche technische Herausforderungen erschwert, welche von Azvine et al. (2005) in drei Ebenen unterteilt werden, die in Kapitel 4.2 als drei von vier Hauptkategorien in Erscheinung treten.

**Analytics.** Traditionelle BI-Lösungen verlangen, dass ein Experte mit hohem Fachwissen für die Konfiguration und Durchführung von Analysen verantwortlich ist (Hang und Fong 2010). Diese manuelle Interaktion hat jedoch eine Latenz bei der Datenanalyse zur Folge. Zur Minimierung dieser Latenz müssen Analysewerkzeuge in RTBI-Lösungen ein hohes Maß an Automatisierung besitzen. Die Analysesoftware muss aus verschiedenen Methoden wählen und diese automatisch anwenden können. Weichen die Ergebnisse von vorher festgelegten Nutzeranforderungen ab, so werden sie mit anderen Parametern wiederholt. Die dabei entstehenden Analyseergebnisse werden anschließend in „Report-Form“ präsentiert, welche es dem Nutzer erlauben entsprechende Handlungsalternativen abzuleiten (Azvine et al. 2005).

**Data Integration.** Die Ebene der Data Integration umfasst alle Komponenten, welche für die Bereitstellung geeigneter Daten für die darauf aufbauende Analytics-Ebene verantwortlich sind. Die durchzuführenden Analysen benötigen dabei eine angemessene Datenqualität (Azvine et al. 2005), da nur so aussagekräftige Ergebnisse erzielt werden können. Dies stellt eine Herausforderung für Unternehmen dar, da Daten aus unterschiedlichen Vorkomplexen mit teils sehr verschiedenen, proprietären Strukturen integriert werden müssen. Große Datenvolumen sowie die

möglichst geringe Integrationszeit stellen weitere Herausforderungen dar. Im Gegensatz zu traditionellen BI-Lösungen muss ein Update des Datenbestandes bei RTBI-Lösungen in kürzeren Zeitabständen erfolgen, damit eine hohe Aktualität der Daten gewährleistet werden kann (Chaudhuri et al. 2011). Eine weitere Herausforderung besteht in der parallelen Durchführung von Analysen und Updates auf dem gleichen Datenbestand (Kemper und Neumann 2011).

**Operational.** Die Operationale Ebene umfasst zwei Hauptfunktionen: das „Business Activity Monitoring“, welches der Überwachung von Geschäftsprozessen dient, und die darauf basierende, automatisierte Anpassung dieser Prozesse. Anhand festgelegter betrieblicher Kennzahlen, sogenannten „Key Performance Indicators“, erfolgt eine Messung der Prozessperformance. So können auftretende Abweichungen und Probleme schnell erkannt und entsprechende Gegenmaßnahmen eingeleitet werden. Aus Gründen der Aktualität muss, im Gegensatz zu traditionellen BI-Lösungen, die Anpassung dieser Prozesse in RTBI-Systemen ausnahmslos automatisiert geschehen (Hang und Fong 2010).

Das Hauptmerkmal von RTBI-Lösungen bildet die Minimierung der Latenz zwischen der Erhebung der Daten und der Durchführung einer daraus abgeleiteten Handlung (Seufert und Schiefer 2005; Azvine et al. 2006). Diese Zeitspanne wird als „Action Distance“ bezeichnet (Olsson und Janiesch 2015). Es werden dabei drei Arten der Latenz unterschieden, welche die Verarbeitung der Daten in Echtzeit negativ beeinflussen: (1) Der Zeitraum zwischen dem Anfallen der Daten und der Speicherung im Data-Warehouse wird als Datenlatenz bezeichnet. (2) Die Analyselatenz beschreibt die Zeitspanne von der Speicherung bis hin zur Analyse. (3) Aufbauend auf den Analyseergebnissen muss eine Entscheidung getroffen und anschließend durchgeführt werden – die sogenannte Entscheidungslatenz (Kemper et al. 2010). Während die Minimierung der Daten- und der Analyselatenz primär durch technische Lösungen erfolgt, müssen für die Reduzierung der Entscheidungslatenz auch die Anwender mit einbezogen werden (Tank 2015). Diese organisatorischen Aspekte treten in Kapitel 4.2 als vierte Hauptkategorie in Erscheinung.

Mit dem Ziel der Minimierung der Latenz des Prozesses von der Datenerhebung bis hin zur Durchführung der abgeleiteten Handlungsempfehlung stellen aktuelle Daten somit eine Grundvoraussetzung für das zeitnahe Treffen von Entscheidungen dar. Um aufbauend auf einer BI-eine RTBI-Lösung zu erstellen, gilt es somit zu untersuchen, in welchen Komponenten der zugrundeliegenden Data-Warehouse-Umgebung welche Latenz auftritt und durch welche technischen bzw. organisatorischen Maßnahmen diese minimiert werden kann (McKenna 2011).

### 3 Literaturrecherche

Zur Erfassung der aktuellen Forschungsbeiträge im Bereich der RTBI wurde eine systematische Literaturrecherche durchgeführt. Dies geschieht unter Nutzung der weitverbreiteten Methodik von Webster und Watson (2002). Dabei werden zwei grundlegende Ziele verfolgt: (1) die Untersuchung des momentanen Forschungsbereichs von RTBI und (2) die Ableitung einer Forschungsagenda.

Zur Konzeptualisierung des Themas sowie zur Identifikation relevanter Suchbegriffe wurde eine explorative Suche in Literaturdatenbanken (Google Scholar, CiteSeer und Springer Link) durchgeführt. Dies führte zu einer Sammlung von verschiedenen RTBI-bezogenen Suchbegriffen wie „real-time BI“, „real-time data-warehouse“ und „real-time data-warehousing“, welche oft als mögliche Synonyme zu „real-time business intelligence“ verwendet werden. Der betrachtete Publikationszeitraum wurde auf die Jahre 2000 bis 2015 begrenzt, da sich speziell in diesem Zeitraum das zu verarbeitende Datenvolumen vervielfacht hat und datenzentrierte Anwendungen

eine signifikant größere Rolle spielen als davor (Davenport und Harris 2007). Der Fokus der Literaturrecherche lag dabei sowohl auf hochrangigen als auch domänenspezifischen Fachzeitschriften und führenden Konferenzen:

- Fachzeitschriften der AIS Senior Scholar's Basket: „European Journal of Information Systems“, „Information Systems Journal“, „Information Systems Research“ und „Journal of Information Technology“
- BI-spezifische Fachzeitschriften: „Decision Support Systems“, „International Journal of Business Intelligence Research“ und „Business Intelligence Journal“
- Führende Konferenzen: „International Conference on Information Systems“, „European Conference on Information Systems“, „Hawaii International Conference on System Sciences“, „Pacific Asia Conference on Information Systems“ und „Americas Conference on Information Systems“

Wie von Webster und Watson (2002) vorgeschlagen, wurde zunächst durch Kombination der Suchbegriffe eine Keyword-Suche durchgeführt. Daraufhin wurde durch Analyse des Titels und des Abstracts eine engere Auswahl getroffen. Die dabei als relevant eingestuften Arbeiten wurden anschließend im Zuge einer Volltext-Analyse nochmals genauer betrachtet, um so die Eingrenzung weiter zu fokussieren. Zur Vervollständigung der Literatur wurde anschließend eine Vorwärts bzw. Rückwärtssuche durchgeführt. Nach Abschluss der Literaturrecherche wurden insgesamt 33 Publikationen, deren Fokus speziell auf dem Konzept der „real-time“ liegt, als thematisch relevant eingestuft. Tabelle 1 fasst diese Resultate zusammen. Die einzelnen Beiträge können in Kapitel 4 entsprechend ihrer Zuordnungen eingesehen werden.

	Real-Time BI	Real-Time Business Intelligence	Real-Time Data Warehouse	Real-Time Data-Warehousing	Total
Keyword-Suche	774				
Abstract- & Titel-Analyse	39	65	32	30	<u>166</u>
Volltext-Analyse	6	8	4	4	<u>22</u>
Rückwärtssuche	2	2	1	2	<u>7</u>
Vorwärtssuche	1	2	0	1	<u>4</u>
Beiträge je Kategorie	9	12	5	7	<u>33</u>

**Tabelle 1: Ergebnisse der Literatursuche**

Die Keyword-Suche nach den oben aufgeführten Suchbegriffen hatte insgesamt 774 Treffer zur Folge. Davon wurden diejenigen 166 Arbeiten weiter untersucht, die gemäß Titel und Abstract explizit ins Suchfeld passten. Eine anschließend durchgeführte Volltextanalyse diente der Eliminierung zu generischer Arbeiten. Durch Rückwärts- bzw. Vorwärtssuche konnten elf weitere Papers identifiziert und als relevant bewertet werden. Somit besteht der in diesem Paper betrachtete Literaturbestand aus insgesamt 33 Arbeiten, wobei die meisten den Suchbegriffen der „Real-Time Business Intelligence“ (12 Arbeiten, 36%) und der „Real-time BI“ (9, 27%) zuzuordnen sind.

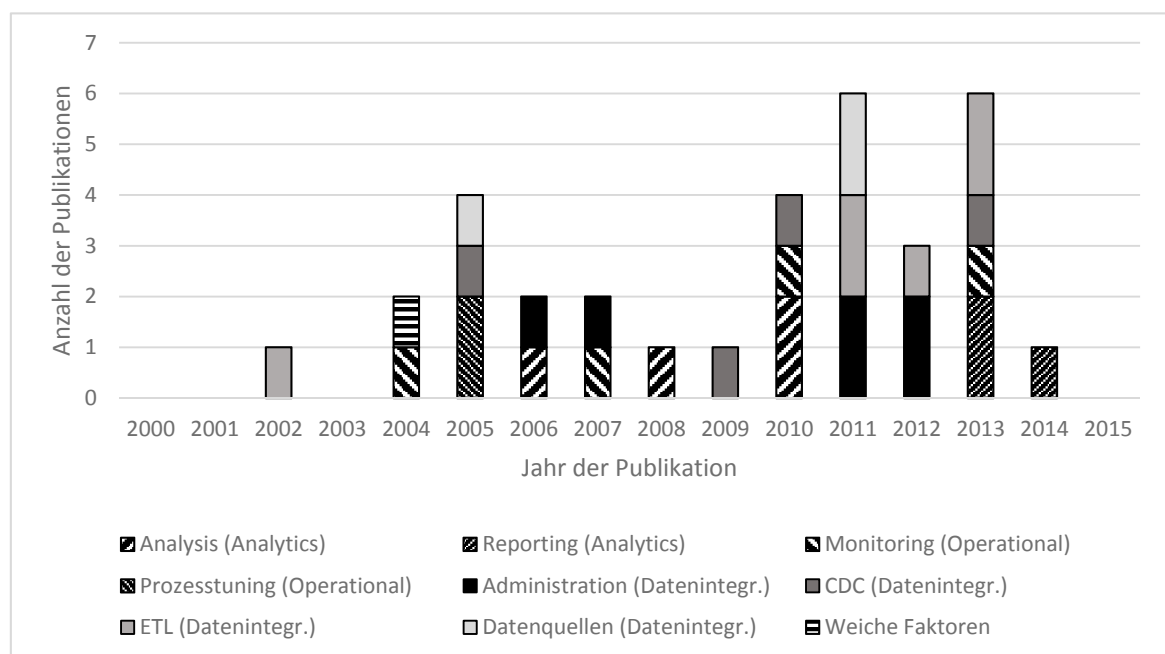
## 4 Präsentation und Analyse der Resultate

In diesem Kapitel werden die Ergebnisse der Literaturrecherche präsentiert und analysiert. Mit Bezug auf die Einteilung von Azvine et al. (2005) werden die 33 identifizierten Arbeiten den

technischen und organisatorischen Herausforderungen von RTBI-Lösungen zugeordnet. Die Analyse erfolgt in zwei Schritten: (1) Zunächst wird die identifizierte Literatur deskriptiv ausgewertet, um den Verlauf und die Forschungsschwerpunkte des wissenschaftlichen Diskurses aufzuzeigen, bevor (2) eine inhaltliche Analyse erfolgt, um den derzeitigen Stand des Wissens zusammenfassend darzustellen.

#### 4.1 Deskriptive Analyse

Im Zuge der deskriptiven Analyse wurden die Arbeiten hinsichtlich ihrer inhaltlichen Schwerpunkte und ihres Publikationsjahres analysiert. Die Schwerpunkte (Analytics, Datenintegration, Operational und weiche Faktoren) werden dabei in Unterkategorien unterteilt, um eine detaillierte thematische Auseinandersetzung zu ermöglichen. Abbildung 1 visualisiert den Zusammenhang zwischen inhaltlichem Schwerpunkt, den jeweiligen Unterkategorien und dem Jahr der Publikation.



**Abbildung 1: Überblick über die behandelten Forschungsschwerpunkte in Abhängigkeit vom Jahr ihrer Publikation**

Abbildung 1 zeigt insgesamt einen Anstieg der Forschungen zum Thema RTBI in den letzten Jahren. Allein seit 2010 sind insgesamt 20 der betrachteten Publikationen veröffentlicht worden.

Insgesamt befassen sich von den 33 Publikationen 21% schwerpunktmäßig mit Fragestellungen, die der Analytischen Ebene zuzuordnen sind. Innerhalb dieser lassen sich zwei Unterkategorien identifizieren: *Analysis* und *Reporting*. Der Bereich der *Analysis* wird von insgesamt vier Beiträgen (12% aller Publikationen), das *Reporting* von drei Arbeiten (9%) thematisiert.

Der Ebene der Datenintegration sind insgesamt 19 Publikationen schwerpunktmäßig zuzuordnen. Dies entspricht rund 57% und damit der absoluten Mehrheit der gesamten untersuchten Literatur. Diese 19 Arbeiten lassen sich weiterhin in vier Schwerpunkte unterteilen: Die Verwaltung der Daten, *Administration*, ist zentraler Bestandteil von sechs Arbeiten (18%). Die Datenintegration unter Nutzung von *Changed Data Capture* (CDC) wird in vier (12%), die Datenintegration mit *ETL*

in sechs Publikationen (18%) behandelt. Die Nutzung potentieller neuer Datenquellen für RTBI-Lösungen wird in drei Arbeiten (9%) thematisiert.

Mit den Aspekten der Operationalen Ebene befassen sich insgesamt sechs Publikationen und damit rund 18% der gesamten Literatur. Zwei Arbeiten können schwerpunktmäßig dem *Monitoring* (6%) und vier Arbeiten dem *Prozesstuning* (12%) zugeordnet werden.

Im Kontext dieser Arbeit konnte nur eine Publikation identifiziert werden, die explizit organisatorische Herausforderungen adressiert und somit in die Kategorie *weiche Faktoren* fällt.

## 4.2 Inhaltliche Analyse

In diesem Abschnitt werden die Arbeiten inhaltlich gruppiert und vorgestellt. Neben den in Kapitel 2 bereits erwähnten vier Hauptkategorien werden die relevanten Publikationen den inhaltlichen Unterkategorien dieser Schwerpunkte zugeordnet.

### 4.2.1 Analytische Ebene

**Analysis.** Während in traditionellen BI-Lösungen Data-Mining-Funktionalitäten für die Datenanalyse genutzt werden, benötigen RTBI-Lösungen eine neue Art von Data-Mining-Verfahren, das Stream Mining. Dies ist durch das kontinuierliche Eintreffen von Datenströmen im System gekennzeichnet, welche anschließend mit Hilfe von Zeitfensterverfahren schnell analysiert und verarbeitet werden können. (Hang und Fong 2010)

Das Finden von Mustern in sehr großen Datenbeständen, wie sie bei RTBI-Lösungen anzutreffen sind, ist eine sehr zeit- und kostenintensive Aufgabe. Aus diesem Grund haben Dass und Mahanti (2006) einen Algorithmus (BDFS(b)-diff-sets) entwickelt, welcher in der Lage ist auch in sehr großen und dichten Datensätzen Muster zu erkennen.

Einen weiteren Forschungsansatz stellt die systematische Integration von Wissensmanagement in den Data-Mining-Prozess für die verbesserte Analyse von Daten in Echtzeit dar. Durch diese Verknüpfung sollen Nutzer schneller in der Lage sein, auf Grundlage von Analyseergebnissen entsprechende Handlungen einzuleiten. Dies führt zu einer Reduzierung der Entscheidungslatenz (You 2010).

Zahlreiche Analysen in RTBI-Lösungen werden periodisch durchgeführt. Dabei werden in der Regel sehr große Datenmengen zusammengeführt und analysiert, was für jede Durchführung einen hohen Zeit- und Rechenaufwand darstellt. Mazur et al. (2008) schlagen zur Lösung dieses Problems den Aufbau eines Repository vor, welches bereits durchgeführte Analysen mitsamt aller Parameter und Ergebnisse speichert. Werden neue Daten in das Data-Warehouse geladen und die Analyse erneut initiiert, so muss nur ein Teil der Ergebnisse basierend auf den neuen Daten aktualisiert werden. Dies hat eine enorme Verbesserung der Analyseperformance zur Folge.

**Reporting.** Trigo et al. (2014) stellen zahlreiche technische Lösungen und praktische Umsetzungen für das Real-Time Reporting im Bereich des Rechnungswesens vor. Dazu zählen unter anderem Cloud Computing, Enterprise Application Integration und Mobile Devices.

Einen weiteren Ansatz bildet der Aufbau einer Reporting-Lösung in Kombination mit einer zugrunde liegenden In-Memory-Datenbank (Santhosh Baboo und Prabhu 2013).

Performance-Dashboards stellen einen wichtigen Bestandteil von BI-Lösungen dar, da sie den Ablauf von Prozessen in Form von KPIs überwachen. Steinkamp und Mühlbauer (2013) stellen „HyDash“ vor, ein webbasiertes Dashboard für RTBI. Hauptbestandteil dieser Lösung ist die



Darstellung von KPIs auf Grundlage vordefinierter SQL-Abfragen sowie ein Interface, um ad-hoc SQL-Abfragen zu erstellen. Dadurch wird es auch Nutzern, welche nicht über Kenntnisse spezifischer Datenbankabfragen verfügen ermöglicht, aussagekräftige Analysen zu erzeugen. Je nach Art der KPIs kann ein Dashboard sowohl auf strategischer, taktischer als auch auf operativer Ebene einer RTBI-Lösung eingesetzt werden.

#### 4.2.2 Datenintegration-Ebene

**Administration.** Die parallele Durchführung von Analysen auf einem sich ständig ändernden Datenbestand ist Bestandteil zahlreicher Forschungsarbeiten. Rutz et al. (2012) zeigen auf, dass durch eine Replikation der Daten und somit der Trennung von Analysen und Updates ein enormer Performancegewinn erzielt werden kann. Eine mögliche technische Umsetzung dieser Replikation stellt die Erstellung eines „Snapshots“ der Datenbank dar. Auf dieser Grundlage können die Analysen durchgeführt werden, während parallel eine Aktualisierung der zugrunde liegenden Daten erfolgen kann (Kemper und Neumann 2011).

Aufgrund ihrer hohen Performance wird durch die Nutzung von In-Memory-Datenbanken ebenfalls die parallele Durchführung von Updates und Analysen auf einem gemeinsamen Datenbestand ermöglicht (Acker et al. 2011).

Eine andere Möglichkeit dies zu erreichen, besteht im Einsatz von verteilten Dateisystemen mit „Master-Nodes“, welche jeweils das Verhalten mehrerer „Slave-Nodes“ koordinieren (Santos et al. 2012). Ebenfalls mit dem Ansatz des verteilten Arbeitens in RTBI-Lösungen befassen sich Chang et al. (2006). Die von ihnen vorgestellte Lösung „BigTable“ ist für die Arbeit mit sehr großen Datenmengen, bis hin zu mehreren Petabytes, optimiert und wird bereits in der Praxis von Anwendungen wie Google Analytics und Google Earth genutzt.

Eine große Herausforderung bei der Bereitstellung qualitativ hochwertiger Daten stellt die Kombination zweier verschiedener Perspektiven dar: die IT-Sicht, bestehend aus allen relevanten technischen Komponenten und Prozessen und die Business-Sicht, repräsentiert durch übergeordnete Geschäftsziele und -strategien. Eine Annäherung dieser beiden Sichten soll durch die Einführung eines gemeinsamen Begriffsverständnisses in Form einer Ontologie erfolgen (Cui et al. 2007). Dadurch können insbesondere im Bereich der Kommunikation, der Interoperabilität und des Wissensmanagements Verbesserungen erzielt werden.

**CDC.** Mehrere Arbeiten befassen sich mit dem Ansatz des Change Data Capture (CDC) für die Integration neuer Daten. Der Grundsatz dieses Verfahrens liegt im Erkennen von Veränderungen im Datenbestand. Brobst und Pareek (2009) stellen allgemeine Ansätze für CDC vor, wohingegen der spezifische Einsatz von CDC bei Organisationen, welche Altsysteme als Datenquelle nutzen, von Oliver (2013) thematisiert wird. Ankorion (2005) betrachtet weiterhin die technischen Herausforderungen beim Einsatz von CDC. Eine Abwandlung davon bildet das „Incremental Continuous Refresh“, welches neue Daten erst dann aus den Vorsystemen in das Data-Warehouse lädt, wenn das System die Notwendigkeit dafür sieht und das Update freigibt (Chen et al. 2010).

**ETL.** Der formale Aufbau des ETL-Prozesses, im Kontext von RTBI, ist Bestandteil mehrerer Publikationen (McKenna 2011; Kakish and Kraft 2012). Bruckner et al. (2002) nutzen in ihrem Artikel die bestehende J2EE-Technologie, ETLets und Enterprise Java Beans zum Aufbau einer ETL-Prozess-Umgebung für die Integration von Daten in Echtzeit.

Eine weitere Möglichkeit für die Optimierung des ETL-Prozesses im Kontext von „real-time“ stellt die Umwandlung in einen ELT-Prozess dar (Freudenreich et al. 2013), wobei die Daten ohne vorherige Änderung oder Normalisierung direkt ins Data-Warehouse geladen werden.

Santos et al. (2011) stellen ein „Middleware-Tool“ vor, welches zwischen den Datenquellen und dem eigentlichen ETL-Prozess agiert. Aufgrund seiner Fähigkeit, neue Datensätze in sehr kurzer Zeit einzufügen, können mögliche Auswirkungen auf die Datenbank-Performance minimiert werden.

Die Bereitstellung qualitativ hochwertiger Daten für spätere Analysen wird durch eine geeignete Verbindung von IT- und Business-Sicht im Unternehmen erschwert. Arocena et al. (2013) stellen ein formales Framework zur Lösung dieses „Vivifications-Problems“ vor.

**Neue Datenquellen.** Sensoren und RFID bilden neue mögliche Datenquellen für „real-time“ Daten. Während Panian (2005) einen Überblick über grundlegende Architekturen gibt und zwei theoretische Ansätze vergleicht, thematisiert Marinos (2005) in seiner Arbeit neben der formalen Architektur mögliche Fragestellungen, welche vor der Nutzung von RFID in RTBI-Lösungen beachtet werden müssen.

Die Integration von Social Media-Inhalten wird von Shroff et al. (2011) mit Hilfe eines „Enterprise Information Fusion Framework“ realisiert.

#### 4.2.3 Operationale Ebene

**Monitoring.** Wang und Wang (2005) schlagen die Einführung eines Software-Agenten-basierten Frameworks zum Monitoring vor. Jeder einzelne Software-Agent nimmt seine Umgebung durch das Erfassen von Ereignissen und dem Status ablaufender Prozesse wahr. Mit Hilfe festgelegter „Business-Rules“ werden auf Grundlage der erfassten Daten entsprechende korrigierende Handlungen initiiert.

Einen Schritt weiter gehen Finnie und Barker (2005) mit der Einführung des Lernverfahrens des fallbasierten Schließens für Software-Agenten. Wird eine Abweichung im Prozess entdeckt, so nutzen die Agenten die Erfahrungen früherer Probleme und Ereignisse für eine möglichst selbstständige Lösung der Problemstellung.

**Prozesstuning.** Zentraler Bestandteil der Operationalen Ebene ist die Überwachung der Erreichung vorher definierter Geschäfts- und Prozessziele. Golfarelli et al. (2004) schlagen den Aufbau einer kompletten Business-Process-Management-Architektur vor, welche nicht nur den reinen Data-Warehouse-Prozess, sondern alle Komponenten einer BI-Lösung integriert. Gao und Xu (2010) beschreiben ein Framework, das die automatisierte Abwicklung von „Exceptions“ beinhaltet, welche den erfolgreichen Abschluss eines Prozesses verhindern.

Weiterhin existieren verschiedene Möglichkeiten um den effizienten Betrieb einer RTBI-Lösung sicherzustellen. Die Minimierung operationeller Risiken durch die Einführung eines Organisation „Risk Managements“ wird von Azvine et al. (2007) empfohlen. Eine weitere Option besteht im Einsatz des ACP-Ansatzes (Artificial systems/societies, Computational experiments and Parallel execution). Durch den parallelen Betrieb eines zweiten, identischen Systems können verschiedene Problemstellungen simuliert und so wichtige Erkenntnisse für den Betrieb des eigentlichen Systems gewonnen werden. (Li und Shen 2013)

#### 4.2.4 Organisatorische Aspekte

Neben den technischen Faktoren spielen auch organisatorische Aspekte eine wesentliche Rolle, welche gleichermaßen für den erfolgreichen Betrieb von RTBI-Lösungen berücksichtigt werden müssen. Dazu zählen sowohl kulturelle als auch organisatorische Herausforderungen, mit denen sich Nutzer einer RTBI-Lösung tagtäglich konfrontiert sehen. Eckerson (2004) legt in seiner Arbeit dar, dass die Akzeptanz der Nutzer ein kritisches Erfolgskriterium bei der Einführung einer RTBI-Lösung darstellt. Nur wenn sich Nutzer über die Vor- aber auch über die möglichen Nachteile, wie z.B. einer geringeren Datenkonsistenz von Echtzeit-Daten bewusst sind, wird die erfolgreiche Arbeit mit RTBI-Lösungen in der Praxis ermöglicht.

### 5 Aktueller Forschungsbedarf

Im Zuge der inhaltlichen Analyse wurden neben zentralen Forschungsschwerpunkten auch Lücken in der bestehenden Literatur aufgedeckt. Bei Einordnung der als relevant eingestuften Arbeiten in die vier zuvor festgelegten Hauptkategorien wurden drei Themengebiete (Entscheidungsprozesse, weiche Faktoren und komplette Implementierung der Architektur) identifiziert, die nach unseren Erkenntnissen in der RTBI-Forschung bisher nur unzureichend untersucht worden sind. Der hier ermittelte Forschungsbedarf motiviert damit weitergehende Untersuchungen. Zusätzlich zu vielversprechenden Forschungsgebieten werden in diesem Kapitel auch geeignete Methoden zur Untersuchung aufkommender Forschungsfragen vorgestellt.

**Entscheidungsprozesse.** RTBI-Lösungen versorgen den Nutzer mit aufbereiteten Daten in einer organisierten Form. Jedoch sind viele dieser Systeme nicht in der Lage automatisiert Entscheidungen zu treffen. Die Ableitung von Handlungen auf Grundlage der analysierten Daten stellt somit eine große Herausforderung bei der Arbeit mit RTBI-Lösungen dar. Die Reduktion der Entscheidungslatenz wird in der bestehenden Literatur nur selten behandelt. Hier kann die Entwicklung und Evaluation von automatisierten Agenten, welche selbstständig die Ausführung von Prozessen initiieren, einen vielversprechenden Forschungsansatz darstellen.

**Weiche Faktoren.** Organisatorische Herausforderungen sind momentan weitestgehend unerforscht (Eckerson 2004). Nur sehr wenige Publikationen thematisieren den Einfluss von fachlichen oder technischen Anwendern auf RTBI-Projekte. Eckerson (2004) sieht die starke Einbeziehung von Geschäftseinheiten, die klare Formulierung von Zielstellungen und ein hohes Verständnis der zugrunde liegenden Geschäftsprozesse als Schlüsselfaktoren für die erfolgreiche Implementierung von RTBI-Projekten. Bestandteil zukünftiger Forschungen könnte somit die Identifikation weiterer relevanter weicher Faktoren sein. Weiterhin ist es wichtig Fachexperten, welche von der Umstellung einer BI- zu einer RTBI-Lösung betroffen sind, in diesen Prozess aktiv einzubeziehen. Nur wenn es gelingt, mögliche Vorbehalte dieser Nutzer abzubauen, wird eine erfolgreiche Implementierung ermöglicht (Eckerson 2004). Wie bereits in bisherigen BI-Forschungsprojekten gezeigt, bilden Interviews mit Fachexperten in Kombination mit Beobachtungen aus Feldstudien einen adäquaten Ansatzpunkt zur Bewältigung dieser Herausforderungen.

**Komplette Implementierung der Architektur.** RTBI-Lösungen werden oft als eine Vielzahl an Komponenten beschrieben, deren erfolgreiche Integration große Probleme bereitet (Sahay and Ranjan 2008). Organisationen benötigen ein konsistentes, kostengünstiges und benutzerfreundliches BI-System mit geringen Wartungskosten, um sowohl den Bedürfnissen der Anwender aus fachlichen Kontexten als auch der aus technischen gerecht zu werden. Bei einer Vielzahl der untersuchten Publikationen lag der Fokus jedoch auf spezifischen RTBI-Komponenten, ohne

Rücksicht auf mögliche Probleme bei der Integration mit anderen Komponenten zu nehmen. Zusätzlich wurde deren Umsetzung und Implementierung oft gar nicht bzw. nur unter Laborbedingungen realisiert. Aus diesen Gründen stellt die Untersuchung und Implementierung des kompletten BI-Prozesses, ausgehend von den Datenquellen bis hin zur Durchführung von Analysen und dem Erstellen von „Reports“ unter echten Bedingungen, einen vielversprechenden Forschungsansatz dar.

## 6 Konklusion und Ausblick

Das Konzept der RTBI hat in den letzten Jahren zunehmend an Relevanz gewonnen und ist momentan Bestandteil zahlreicher Forschungsarbeiten. Der Forschungsschwerpunkt liegt dabei häufig auf der Betrachtung spezifischer technischer Problemstellungen und entsprechender Lösungen. Durch die Fokussierung vieler RTBI-Publikationen auf die Umsetzung spezieller Verfahren werden eine systematische Methodologie und die Definition einer Forschungsagenda notwendig. Diese Arbeit adressiert beide Problemstellungen mit Hilfe einer systematischen Literaturrecherche.

Die Ergebnisse der Untersuchung zeigen, dass insbesondere im technischen Bereich zahlreiche Publikationen existieren, denen eine gemeinsame und einheitliche Methodologie fehlt. Weiterhin mangelt es an Forschungsarbeiten, welche die Interaktion mit und die Verantwortung von Anwendern im Kontext der eingesetzten RTBI-Lösungen untersuchen. Eine Erweiterung der bisher überwiegend technischen BI-Perspektive um „weiche“ Faktoren sollte somit einen Schwerpunkt künftiger Untersuchungen darstellen. Weiterhin sollten zukünftige Forschungsarbeiten ganzheitliche RTBI-Lösungen in der Praxis implementieren und deren Einsatz unter realen Bedingungen testen, um den praktischen Nutzen des „real-time“-Paradigmas und seiner Effekte zu erforschen.

Folgende Limitationen der Arbeit sind zu beachten. Der Fokus der Arbeit wurde bewusst auf „real-time“ gelegt. Artverwandte Konzepte wie „Operational Business Intelligence“, „Process-centric Business Intelligence“ oder „Business Process Intelligence“ wurden dementsprechend nicht betrachtet. Zukünftige Arbeiten könnten diese Literaturbereiche ergänzen.

## 7 Literaturverzeichnis

- Acker O, Gröne F, Blockus A, Bange C (2011) In-memory analytics – strategies for real-time CRM. *J Database Mark Cust Strateg Manag* 18:129–136. doi: 10.1057/dbm.2011.11
- Ankorion I (2005) Change Data Capture Efficient ETL for Real-Time BI. *DM Rev* 15:36–43.
- Arocena P, Miller R, Mylopoulos J (2013) The Vivification Problem in Real-Time Business Intelligence: A Vision. In: *Enabling Real-Time Business Intelligence*. Springer Berlin Heidelberg, pp 37–49
- Azvine B, Cui Z, Majeed B, Spott M (2007) Operational risk management with real-time business intelligence. In: *BT Technology Journal*. Springer Berlin Heidelberg, pp 154–167
- Azvine B, Cui Z, Nauck DD (2005) Towards real-time business intelligence. *BT Technol J* 23:214–225. doi: 10.1007/s10550-005-0043-0
- Azvine B, Cui Z, Nauck DD, Majeed B (2006) Real Time Business Intelligence for the Adaptive Enterprise. In: *E-Commerce Technology, 2006. The 8th IEEE International Conference on and Enterprise Computing, E-Commerce, and E-Services, The 3rd IEEE International Conference on*. pp 1–11

- Brobst S, Pareek A (2009) New Trends in Data Acquisition Services for the Real-Time Enterprise. *Bus Intell J* 14:52–58.
- Bruckner RM, List B, Schiefer J (2002) Striving towards Near Real-Time Data Integration for Data Warehouses. *DaWak* 317–326.
- Chang F, Dean J, Ghemawat S, et al (2006) Bigtable: A distributed storage system for structured data. In: 7th Symposium on Operating Systems Design and Implementation (OSDI '06), November 6-8, Seattle, WA, USA. pp 205–218
- Chaudhuri S, Dayal U, Narasayya V (2011) An overview of business intelligence technology. *Commun ACM* 54:88–98. doi: 10.1145/1978542.1978562
- Chen LCL, Rahayu W, Taniar D (2010) Towards Near Real-Time Data Warehousing. In: *Advanced Information Networking and Applications (AINA), 2010 24th IEEE International Conference on*. pp 1150–1157
- Cui Z, Damiani E, Leida M (2007) Benefits of Ontologies in Real Time Data Access. In: *2007 Inaugural IEEE International Conference on Digital Ecosystems and Technologies (IEEE DEST 2007)*. pp 392–397
- Dass R, Mahanti A (2006) An Efficient Algorithm for Real-Time Frequent Pattern Mining for Real-Time Business Intelligence Analytics. In: *Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06)*. pp 1–10
- Davenport TH, Harris JG (2007) *Competing on analytics: The new science of winning*. Harvard Business Review Press
- Eckerson W (2004) The “Soft Side” of Real-Time BI. *DM Rev* 14:30–33.
- Finnie G, Barker J (2005) Real-time business intelligence in multi-agent adaptive supply networks. In: *Proceedings - 2005 IEEE International Conference on e-Technology, e-Commerce and e-Service, EEE-05*. pp 218–221
- Freudenreich T, Furtado P, Koncilia C, et al (2013) An On-Demand ELT Architecture for Real-Time BI. In: *Business Intelligence for the Real Time Enterprise (BIRTE 2012)*. Springer Berlin Heidelberg, pp 50–59
- Gao S, Xu D (2010) Real-Time Exception Management Decision Model (RTEMMDM): Applications in intelligent agent-assisted decision support in logistics and anti-money laundering domains. In: *Proceedings of the Annual Hawaii International Conference on System Sciences (HICSS'10)*. pp 1–10
- Geerdink B (2013) A Reference Architecture for Big Data Solutions. In: *The 8th International Conference for Internet Technology and Secured Transactions*. pp 71–76
- Gluchowski P, Gabriel R, Dittmar C (2008) *Management Support Systeme und Business Intelligence*. Springer Berlin Heidelberg
- Golfarelli M, Rizzi S, Cella I (2004) Beyond Data Warehousing: What's Next in Business Intelligence? In: *Proceedings of the 7th ACM international workshop on Data warehousing and OLAP*. pp 1–6
- Hang Y, Fong S (2010) Real-time business intelligence system architecture with stream mining. *2010 5th Int Conf Digit Inf Manag ICDIM 2010* 29–34. doi: 10.1109/ICDIM.2010.5664637
- Kakish K, Kraft T a (2012) ETL Evolution for Real-Time Data Warehousing. In: *Proceedings of the Conference on Information Systems Applied Research*. pp 1–12
- Kemper A, Neumann T (2011) One Size Fits all , Again! The Architecture of the Hybrid OLTP & OLAP Database Management System HyPer. In: *Enabling Real-Time Business Intelligence*. Springer Berlin Heidelberg, pp 7–23
- Kemper H-G, Baars H, Mehanna W (2010) *Business Intelligence - Grundlagen und praktische Anwendung*. Vieweg+Teubner
- Li J, Shen Z (2013) A real-time business intelligence system based on the ACP approach. In: *Proceedings of 2013 IEEE International Conference on Service Operations and Logistics, and Informatics, SOLI 2013*. pp 345–349

- M. Tank D (2015) Enable Better and Timelier Decision-Making Using Real-Time Business Intelligence System. *Int J Inf Eng Electron Bus* 7:43–48. doi: 10.5815/ijieeb.2015.01.06
- Marinos G (2005) RFID and the Value of Context: How A Disruptive Technology Promises Real-Time Intelligence. *DM Rev* 15:38–40.
- Mazur P, Murlewski J, Kaminski M, et al (2008) Universal Database Analysis Repository Model for Use in Near Real-Time Database Warehouse Environment. In: *Modern Problems of Radio Engineering, Telecommunications and Computer Science*. Lviv-Slavsko, Ukraine, pp 577–579
- McKenna JP (2011) Moving Toward Real-Time Data Warehousing. *Bus Intell J Vol* 18, No3 16:14–20.
- Oliver F (2013) Real-Time Data Warehousing for Legacy Applications. *Bus Intell J* 18:29–35.
- Olsson L, Janiesch C (2015) Real-time Business Intelligence und Action Distance: Ein konzeptionelles Framework zur Auswahl von BI-Software. In: *Wirtschaftsinformatik Proceedings*. pp 691–705
- Panian Ž (2005) Real-time Business Intelligence Opportunities Derived from Radio Frequency Identification Data. In: *4th WSEAS International Conference on System Science and Simulation in Engineering*. pp 21–27
- Rutz D, Nelakanti T, Rahman N (2012) Practical Implications of Real Time Business Intelligence. *J Comput Inf Technol* 20:257–264.
- Sahay BS, Ranjan J (2008) Real time business intelligence in supply chain analytics. *Inf Manag Comput Secur* 16:28–48. doi: 10.1108/09685220810862733
- Santhosh Baboo S, Prabhu J (2013) Designing Real Time Business Intelligence Reporting for Better Business Insights. *Int J Comput Appl* 76:12–17.
- Santos RJ, Bernardino J, Vieira M (2012) Leveraging 24/7 Availability and Performance for Distributed Real-Time Data Warehouses. In: *Computer Software and Applications Conference (COMPSAC), 2012 IEEE 36th Annual*. pp 654–659
- Santos RJ, Bernardino J, Vieira M (2011) 24/7 real-time data warehousing: A tool for continuous actionable knowledge. In: *Computer Software and Applications Conference (COMPSAC), 2011 IEEE 35th Annual*. pp 279–288
- Seufert A, Schiefer J (2005) Enhanced Business Intelligence - Supporting business processes with real-time business analytics. In: *Proceedings - International Workshop on Database and Expert Systems Applications, DEXA*. pp 919–925
- Shroff G, Agarwal P, Dey L (2011) Enterprise information fusion for real-time business intelligence. In: *Information Fusion (FUSION), 2011 Proceedings of the 14th International Conference on*. pp 1–8
- Stalk G (1988) Time- The Next Source of Competitive Advantage. *Harv Bus Rev* Volume: 65:Pages: 41–51.
- Steinkamp M, Mühlbauer T (2013) HyDash: A Dashboard for Real-Time Business Intelligence based on the HyPer Main Memory Database System. In: *Studierendenprogramm der GI-Fachtagung Datenbanksysteme für Business, Technologie und Web (BTW 2013)*. pp 217–226
- Trigo A, Belfo F, Estébanez RP (2014) Accounting Information Systems: The Challenge of the Real-time Reporting. *Procedia Technol* 16:118–127. doi: 10.1016/j.protcy.2014.10.075
- Wang M, Wang H (2005) Intelligent Agent Supported Business Process Management. In: *Proceedings of the 38th Annual Hawaii International Conference on (HICCS'2005)*. pp 1–10
- Webster J, Watson RT (2002) Analyzing the past to prepare for the future: Writing a literature review. *MIS Q Vol* 26 No 2 26:xiii–xxiii. doi: 10.1.1.104.6570
- You H (2010) A Knowledge Management Approach for Real-Time Business Intelligence. In: *Intelligent Systems and Applications (ISA), 2010 2nd International Workshop on*. pp 1–4

# Systemarchitektur eines mobilen Empfehlungssystems mit Echtzeitanalysen von Sensordaten für Asthmatiker

Nikolai Bock<sup>1</sup>, Matthias Scholz<sup>2</sup>, Gunther Piller<sup>2</sup>, Klaus Böhm<sup>3</sup>, Hartmut Müller<sup>1</sup>, Dennis Fenchel<sup>3</sup>, Torsten Sehlinger<sup>4</sup>, Martin van Wickeren<sup>3</sup> und Wolfram Wiegers<sup>3</sup>

<sup>1</sup> Hochschule Mainz, i3mainz, [firstname.lastname@hs-mainz.de](mailto:firstname.lastname@hs-mainz.de)

<sup>2</sup> Hochschule Mainz, Fachbereich Wirtschaft, [firstname.lastname@hs-mainz.de](mailto:firstname.lastname@hs-mainz.de)

<sup>3</sup> health & media GmbH, Darmstadt, [firstname.lastname@health-media.de](mailto:firstname.lastname@health-media.de)

<sup>4</sup> Bluestone Technologies GmbH, Wörrstadt, [firstname.lastname@bluestone-tec.com](mailto:firstname.lastname@bluestone-tec.com)

## Abstract

Das Projekt *ActOnAir* verfolgt einen neuartigen Ansatz zur Erfassung, Kombination und Analyse persönlicher Belastungsinformationen und feinmaschiger Umweltdaten für Asthmatiker. Individuelle Handlungsalternativen werden identifiziert, bewertet und überwacht. Dieser Beitrag stellt Gestaltungsziele und Kernanforderungen für die zu entwickelnden Softwarekomponenten vor und beschreibt eine hierzu passende Systemarchitektur sowie deren prototypische Umsetzung.

## 1 Einleitung und Motivation

Luftverschmutzung stellt das größte auf Umweltfaktoren basierende Gesundheitsrisiko dar. So besagen statistischen Hochrechnungen der Weltgesundheitsorganisation, dass jeder achte Todesfall weltweit auf verschmutzte Luft zurückgeht. Auch deutsche Städte überschritten 2013 die von der EU festgesetzten Grenzwerte für gesundheitsgefährdende Stoffe, wie Feinstaub als auch Stickstoffdioxid, und damit auch die strengereren Leitwerte, welche die WHO zum Gesundheitsschutz empfiehlt (UBA 2014). Menschen mit Asthma reagieren besonders sensitiv gegenüber partikulären Luftverunreinigungen. Obwohl der Einfluss von Umweltfaktoren auf das Auftreten von Asthmasymptomen vielfach erforscht wird (z.B. Schachter et al. 2015) und die Menge der verfügbaren Informationen über die Umwelt stark anwächst, ist es schwierig, Asthmapatienten individuelle Handlungsempfehlungen für ihren Alltag zur Verfügung zu stellen. Maßgebliche Gründe hierfür sind:

- Eine umfassende und engmaschige Erfassung der Beschwerdesymptome von Individualpersonen und der auf sie einwirkenden Umweltfaktoren, findet nicht statt. Viele Schadstoff- und Umweltdaten werden nur an wenigen Orten gemessen und sind Einzelpersonen schwer zuzuordnen.
- Personenbezogene Analysen und situative Vorhersagen fehlen. Aufgrund unterschiedlicher Disposition ist der Einfluss von Schadstoffschwellwerten, weiterer Umweltfaktoren sowie der

situativen und persönlichen Konstitution für ein Auftreten von Belastungssymptomen sehr individuell.

Die Menge der verfügbaren Informationen über die Gesundheitswerte von Individuen und der für sie maßgeblichen Umweltfaktoren, wird durch eine wachsende Anzahl tragbarer Sensoren und öffentlich zugänglicher Umweltdatenbanken stark anwachsen. Umso wichtiger ist es, aus diesen Daten auch verlässliche und situationsgerechte Handlungsempfehlungen für betroffene Personen abzuleiten.

Der beschriebenen Problematik soll mit einem neuen Hard- und Softwaresystem maßgeblich Abhilfe geschaffen werden. Ein entsprechendes Vorhaben startete Anfang 2015. Dieser Beitrag beschreibt die Zielsetzung und Anforderungen sowie erste Ergebnisse. Der Schwerpunkt liegt auf der Architektur und einer ersten prototypischen Implementierung des Softwaresystems. Das Vorgehen basiert auf dem gestaltungsorientierten Ansatz von Hevner et al. (2004).

## 2 Stand der Dinge

Seit ca. 10 Jahren gibt es zunehmend Projekte zur Messung lokaler Umweltbelastungen, insbesondere zur Luftqualität (siehe z.B. Copenhagen Wheel (2014)). Für persönliche Belastungswerte und Gesundheitsdaten werden neue Sensor-Messkonzepte, wie *People as Sensors*, immer wichtiger (z.B. Sagl et al. 2015). Auch der Trend *Quantified Self*, u.a. mit kommerziellen Plattformen wie Apple Health und Google Fit, spielt eine immer größere Rolle. Speziell die Integration von grundlegend verschiedenen Messverfahren und Möglichkeiten einer flexiblen Erweiterung zur Integration neuer Sensoren und ihrer Daten, stellen noch weitestgehend offene Herausforderungen dar, die im Rahmen dieses Projektes für die Domäne Atemwegsbelastungen/ Asthma adressiert werden.

Für personenbezogene Analysen im Gesundheitsbereich wird Data Mining von Sensordaten immer wichtiger (z.B. Sow et al. 2013). Die Entwicklung von Verfahren, welche die dynamische Natur aller möglichen Einflussfaktoren berücksichtigt, sowie die Interpretierbarkeit der Ergebnisse gewährleistet, ist eine wesentliche methodische Herausforderung. Für Asthma-Monitoring stellt das auf Sequential Pattern Mining beruhende Verfahren von Lee et al. (2010) einen guten Ausgangspunkt dar. Noch nicht ausreichend adressiert wurden jedoch Ansätze zur systematischen Optimierung der verwendeten Mining-Verfahren zum Zweck einer besseren Personalisierung der Ergebnisse, sowie Möglichkeiten zur Bereitstellung von Prognosen und Handlungsempfehlungen in Echtzeit. Beide Aspekte sind Ziel dieser Arbeit.

## 3 Ziele und Anforderungen

Ergebnisartefakt des Projekts ist das zu entwickelnde System, welches die oben beschriebenen Defizite überwindet. Die entsprechenden Gestaltungsziele sind:

- Z1: Die Belastungssituation von Individuen wird ganzheitlich und detailliert erfasst. Alle wesentlichen auf eine Person einwirkenden Luftschadstoffe und Umweltfaktoren werden in kurzen Zeitabständen aufgenommen und mit gleichzeitig aufgezeichneten persönlichen Belastungsinformationen verknüpft.
- Z2: Auf Einzelpersonen zugeschnittene Handlungsempfehlungen werden ohne Zeitverzug bereitgestellt. Verhaltensvorschläge basieren auf aktuellen Umweltdaten und berücksichtigen



die persönliche Disposition, individuelle Beschwerdemuster sowie die augenblickliche Belastungssituation von Betroffenen.

Aus diesen Zielen können Anforderungen an das zu entwickelnde IT-System abgeleitet werden. Da der Schwerpunkt dieses Beitrags die Architektur des Softwaresystems ist, fokussiert er auf Anforderungen, die das Zusammenwirken von Kernkomponenten beschreiben und einen direkten Bezug zu den Gestaltungszielen haben.

Das Gesamtsystem besteht im Wesentlichen aus den folgenden logischen Bausteinen: Einer mobilen Sensorbox, einer Komponente zur Integration und dem Management von Sensordaten (SD), einem Baustein zur Analyse von Sensordaten mit Data Mining-Verfahren (DM) und einer Smartphone-Anwendung für Endanwender (MA). Input für das hier im Mittelpunkt stehende Softwaresystem sind kontinuierlich aufgezeichnete Messdaten für unterschiedliche Luftschadstoffe und Umweltfaktoren sowie Informationen über persönliche Gesundheitswerte. Für deren Erfassung wird eine mobile Einheit (mobile Sensorbox) mit vorhandenen und neu zu entwickelnden Sensoren zur Verfügung gestellt, die in diesem Beitrag nicht weiter beschrieben wird. Zudem werden Daten aus dem Bereich *Quantified Self* genutzt und Informationen über das persönliche Befinden im Kontext von *People as Sensors* mit Hilfe einer mobilen Anwendung abgefragt. Neben diesen, von Einzelpersonen selbst gesammelten Daten, werden auch verfügbare Daten externer Messstationen genutzt. All diese unterschiedlichen Sensordaten sind in der Komponente SD zu integrieren. Die hierfür wesentlichen Anforderungen sind:

- SD1: Die Komponente muss alle Sensordaten integrieren und diese technisch sowie semantisch bereinigen. Insbesondere ist sie in der Lage, Daten von Individualpersonen und externen Messstationen zu kombinieren und auf eine raum-zeitliche Referenz zu projizieren.
- SD2: Die Komponente muss in der Lage sein, Daten neuer Sensoren und zukünftiger externer Services ohne signifikanten Entwicklungsaufwand zu integrieren. D.h. ein Plug & Play-Ansatz muss unterstützt werden.
- SD3: Die Komponente muss es ermöglichen, aktuelle Schadstoff- und Umweltdaten in harmonisierter Form für die Durchführung von Beschwerdeprognosen auf Anfragen von Endanwendersystemen (MA) ohne Zeitverzug bereitzustellen.
- SD4: Die Komponente muss die Fähigkeit haben, harmonisierte Sensordaten kontinuierlich an die Data Mining-Komponente zu übermitteln, sowie diese Daten für mögliche spätere Analyseziele bereitzustellen.

Die Anforderungen SD1 und SD2 beziehen sich primär auf Gestaltungsziel Z1, während SD3 eine Voraussetzung für Z2 ist. SD1, SD2 und SD4 bilden zudem die Grundlage für die Ableitung von Regeln zur Erstellung von Handlungsempfehlungen in der Data Mining-Komponente (DM). In dem hier besprochenen Kontext sind für diese die folgenden Anforderungen wesentlich:

- DM1: Die Komponente muss die Möglichkeit bieten, harmonisierte Messdaten in unterschiedliche Personensegmente zu gliedern und für diese Analyseverfahren durchzuführen. Für erweiterte Untersuchungen müssen auch passende Vergangenheitsdaten von der SD-Komponente abgerufen werden können.
- DM2: Die Komponente muss die Fähigkeit haben, das aus Data Mining-Verfahren gewonnene Regelwerk zur Erstellung von Handlungsempfehlungen für Personensegmente den entsprechenden Endanwendersystemen (MA) bereitzustellen.

- DM3: Das System muss in der Lage sein, aus den von SD kontinuierlich übermittelten Daten ein Verbesserungspotenzial des Regelwerks für Handlungsempfehlungen zu erkennen und entsprechende neue Berechnungen durchzuführen. Ausgangspunkt hierfür können beispielsweise eine höhere Qualität der Datenbasis oder größere Datenvolumen sein.

Die einheitliche Berücksichtigung aller Sensordaten in den verwendeten Mining-Verfahren unterstützt Gestaltungsziel Z1. DM2 und DM3 zielen mit der Personalisierung von Vorhersagen auf Z2 ab. Beispielsweise führt DM3 bei ausreichender Datenbasis zu einer fortschreitenden Differenzierung von Personensegmenten bis hin zur vollständigen Individualisierung der Ergebnisse. DM2 dient zudem der Möglichkeit von Echtzeitempfehlungen aus Z2. Aufbauend auf den Ergebnissen der Data Mining-Komponente muss das Endanwendersystem, d.h. die mobile Anwendung (MA), folgendes leisten:

- MA1: Die Komponente muss in der Lage sein, Handlungsempfehlungen in Echtzeit für den Endanwender bereitzustellen. Input hierfür ist das Regelwerk des DM-Moduls und aktuelle Sensormesswerte aus dem für eine Vorhersage notwendigen Zeitfenster.
- MA2: Das Endanwendersystem muss die Fähigkeit haben, individuelle Handlungsempfehlungen zu erstellen, ohne persönliche Daten in einer Cloud-Umgebung zu persistieren.
- MA3: Die Darstellung der Handlungsempfehlungen muss eine hohe Usability bieten. Dies schließt auch Möglichkeiten einer Personalisierung auf die Bedürfnisse und Wünsche unterschiedlicher Endanwender mit ein.

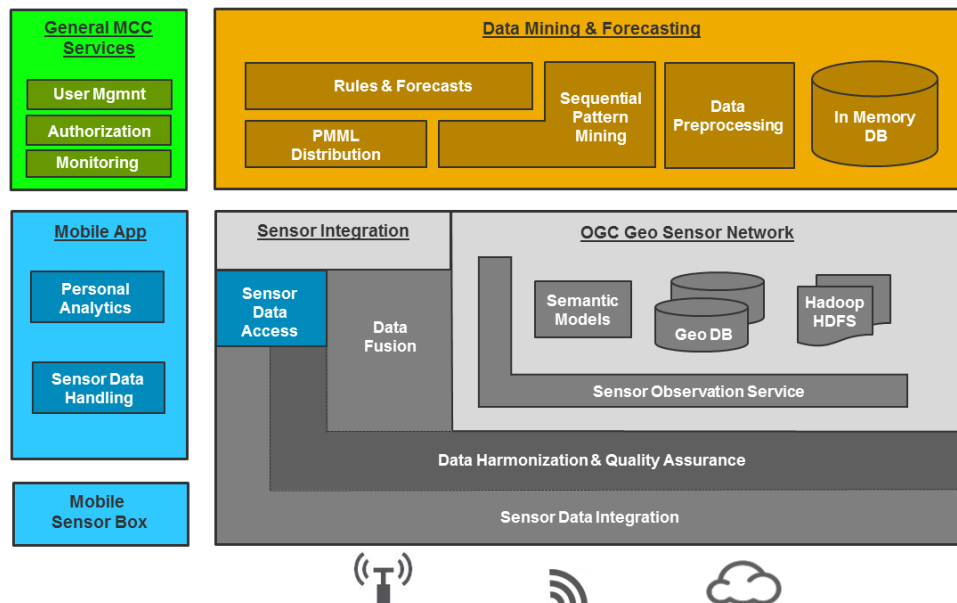
Anforderung MA1 bezieht sich vornehmlich auf Gestaltungsziel Z2. MA2 und MA3 sind maßgebend für die generelle Akzeptanz des Softwaresystems bei Endanwendern.

## 4 Architektur

Eine geeignete Systemarchitektur zur Umsetzung der beschriebenen Projektziele und zentralen Anforderungen zeigt Bild 1. Die drei Hauptkomponenten des Softwaresystems sind:

- Sensorintegration & Geo-Sensornetzwerk (SD): Dieser Baustein ist für die Verarbeitung der heterogenen Sensorinformationen verantwortlich. Die Komponente Sensorintegration bietet Services zur Integration und Kombination von Sensordaten an. Harmonisierte Sensordaten werden im sogenannten Geo-Sensornetzwerk abgelegt. In dem Modul Data Fusion werden Sensordaten von Einzelpersonen mit allgemein verfügbaren Messwerten, z.B. Wetterdaten, mit passendem Raum-Zeit-Bezug aus dem Geo-Sensornetzwerk angereichert und nachfolgend an die Data Mining-Komponente übergeben.
- Data Mining (DM): Hier findet eine Vorverarbeitung von Sensordaten, wie z.B. eine Diskretisierung von Messwerten, statt. Data Mining erfolgt in einem Baustein zur Identifikation häufiger sequenzieller Muster von Sensormesswerten vor Asthmaanfällen und vor beschwerdefreien Tagen. In weiteren Komponenten werden Modelle zur individuellen Vorhersage von Beschwerden berechnet und für mobile Anwendungen bereitgestellt.
- Smartphone-Anwendung (MA): Die mobile Anwendung ist die zentrale Nutzerschnittstelle des Systems. Hier werden individuelle Handlungsempfehlungen mithilfe der in der Data Mining-Komponente berechneten Vorhersagemodelle auf Basis aktueller Sensordaten ausgewertet und dargestellt. Input hierfür können neue, von Patienten erfasste Daten, sowie Umweltdaten aus

dem Geo-Sensornetzwerk sein. Zudem dient die Smartphone-Anwendung als Steuereinheit für alle mobilen und personalisierten Sensoren.



**Bild 1: Systemarchitektur ActOnAir**

Die mobile Sensorbox sowie ein weiterer Baustein mit allgemeinen Diensten, z.B. für Autorisierung und User Management, werden in diesem Beitrag nicht weiter vertieft.

Wie zentrale Anforderungen mithilfe der dargestellten Architektur umgesetzt werden können, wird kurz anhand von Beispielen skizziert: Um individuelle Handlungsempfehlungen in Echtzeit mit aktuellen Sensormesswerten bereitzustellen (MA1), wurde das Analyseverfahren getrennt: Mit Sequential Pattern Mining werden häufige Muster von Sensorwerten, für Tage mit und ohne Asthmabeschwerden, asynchron in der Data Mining-Komponente berechnet. Entsprechend erfolgt hier auch die Berechnung von Vorhersagemodellen. In den mobilen Anwendungen der Endnutzer können diese Modelle jedoch verzögerungsfrei aktuelle Sensordaten auswerten und Empfehlungen geben. Auch die Verbesserung von Vorhersagemodellen, z.B. durch eine stetig wachsende Datenbasis, kann asynchron im Data Mining-Baustein durch selbstlernende Verfahren erfolgen (DM3). Qualitativ höherwertige Modelle werden dann über die PMML-Schnittstelle an Endanwendersysteme für Echtzeitauswertungen verteilt (DM2).

Um einfache Sensordatenerweiterungen (SD2) zu ermöglichen, wurde auf ein Integrationskonzept zurückgegriffen, welches einen Großteil der Messaging Patterns von Hohpe und Woolf (2003) umsetzt. Dadurch können Nachrichten über unterschiedliche Kanäle an Endpunkte, wie z.B. Sensoradapter oder Data Fusion-Prozesse, zur Verarbeitung übergeben werden. Ermöglicht wird auch eine flexible Orchestrierung von Prozessketten, wie beispielsweise zur Harmonisierung, Filterung, Fusion oder Speicherung von Sensordaten. Der Aufwand zur Integration von neuen Sensoren oder Sensormesskonzepten beschränkt sich somit auf die Implementierung neuer Adapter für Sensordatenstrukturen sowie eine gegebenenfalls notwendige Erweiterung von Verarbeitungsprozessen, z.B. durch NLP-Algorithmen. Die Anforderung zur flexiblen Sensorintegration lässt sich so erfolgreich umsetzen (Bock und Böhm 2015).

## 5 Erste prototypische Implementierung

Dieser Abschnitt skizziert ausgewählte Details einer ersten prototypischen Implementierung des ActOnAir-Softwaresystems. Das Sensordatenmanagement besteht aus zwei zentralen Bausteinen, der Sensorintegration und dem Geo-Sensornetzwerk. Für Letzteres wird das SOS-Framework von 52°north (52n SOS 2015) eingesetzt. Dieses implementiert die aktuelle Definition des OGC SOS Standards, sowie Modelle für Informationen zur Luftqualität. Als Datenspeicher wird das DBMS PostgreSQL mit der räumlichen Erweiterung PostGIS verwendet. Das SOS-Framework erlaubt es auch Daten anderer Quellen virtuell einzubinden. Aufbauend auf den OGC-Standards (OGC 2015) bietet das System eine einheitliche API für den produzierenden und konsumierenden Zugang, inkl. räumlich/zeitlicher Filteroperatoren, von Sensorinformationen an. Die Sensorintegration setzt auf Spring XD auf (Spring XD 2015), das u.a. auf Spring Integration und den hierin implementierten Messaging Patterns (Hohpe und Woolf 2003) aufbaut. Für einen ersten Proof of Concept (PoC) wurden Module implementiert, die auf neu entwickelten Adaptern basieren. Sie erlauben eine erste Sensordatenverarbeitung und Kommunikation mit den anderen Bausteinen der Anwendung. Hierzu zählen die Integration von Sensordaten aus der mobilen Anwendung, die Datenbereitstellung an die Data Mining-Komponente sowie das Ablegen der Sensorinformationen durch den Sensor Observation-Service. Zudem wurden erste Adapter für Umweltdaten-Services unterschiedlicher Datenformate (CSV, JSON, NetCDF) entwickelt.

Für die Data Mining-Komponente wurde die In Memory-Plattform SAP HANA verwendet. Die Leistungsmerkmale von HANA sind bei den letztendlich zu erwartenden hohen Datenvolumen insbesondere für initial notwendige explorative Data Mining-Analysen sowie eine performante Verbesserung von Vorhersagemodellen zweckmäßig. Um die Möglichkeiten von HANA auszuschöpfen, wurde das Programmiermodell entsprechend gewählt - z.B. datenbanknahe Umsetzung der Anwendungslogik mit SQLScript, einfaches de-normalisiertes Datenmodell, spaltenorientierte Tabellen, weitgehende Parallelisierung von Prozeduren. Für das Sequential Pattern Mining wurde eine R-Implementierung des SPADE-Algorithmus verwendet (Buchta et al. 2015). Häufige Muster für Messwerte zu persönlichen Belastungen sowie Umweltdaten werden dann als Attribute für die Berechnung von Entscheidungsbäumen genutzt (Lee et al. 2010). Hierzu wurde eine CART-Implementierung von HANA eingesetzt (SAP PAL 2015). Dieses mehrstufige Analyse-Verfahren wurde mit beispielhaften Testdaten zu einigen Asthmasymptomen sowie ausgewählten Wetter- und Luftbelastungswerten erprobt. Das hieraus gewonnene Ergebnis wurde dann als PMML-Dokument an die mobile Anwendung übertragen.

Mit dem hier skizzierten PoC konnte die Integration der Datenflüsse zwischen Sensoren und mobiler Anwendung, sowie den Komponenten Sensorintegration, Geo-Sensornetzwerk und Data Mining beispielhaft realisiert werden. Insbesondere ließ sich die Eignung des Konzepts und deren Umsetzung bezüglich der Anforderungen SD2, SD3, DM1 und DM2 nachweisen.

## 6 Nächste Schritte

Im Zuge der weiteren Entwicklung des Gesamtsystems werden im Bereich Sensordatenmanagement als nächstes Algorithmen zur Harmonisierung und Fusion von Sensordaten entwickelt. Ferner werden die Data Mining-Verfahren mit realistischen Testdaten fachlich evaluiert. Für die sensible Kommunikation mit Endanwendern, einschließlich der Darstellung der entsprechenden Handlungsempfehlungen, wird eine mobile Anwendung konzipiert und entwickelt. Des Weiteren wird ein auf OAuth basierendes Autorisierungskonzept für den gesamten Datenfluss

erprobt. Nach einer abschließenden Integration aller Komponenten wird das Gesamtsystem mit Nutzergruppen und medizinischem Fachpersonal evaluiert. Als Fertigstellungstermin der vollständigen Anwendung ist der Sommer 2016 avisiert.

Gefördert durch das Bundesministerium für Wirtschaft und Energie, ZIM KF; unterstützt durch das HPI Future SOC Lab

## 7 Literatur

- Bock N, Böhm K (2015) Flexible Datenintegration und Migration in ein Sensornetzwerk. AGIT - Journal für Angewandte Geoinformatik 1-2015, Salzburg
- Buchta C, Hahsler M, Diaz D (2015) Package ‘arulesSequences’. <https://cran.r-project.org/web/packages/arulesSequences/arulesSequences.pdf>. Abgerufen am 17.08.2015
- CopenhagenWheel (2014) The Copenhagen Wheel. <http://senseable.mit.edu/copenhagenwheel>. Abgerufen am 17.08.2015
- Hevner AR et al. (2004) Design science in information systems research. MIS Quarterly 28 (1):75-105
- Hohpe G, Woolf B (2003) Enterprise Integration Patterns: Designing, Building, and Deploying Messaging Solutions. Addison-Wesley, Boston
- Lee CH et.al. A Novel Data Mining Mechanism Considering Bio-Signal and Environmental Data with Applications on Asthma Monitoring. Computer Methods and Programs in Biomedicine 101 (1): 44-61
- OGC (2015) OGC Standards and Supporting Documents. <http://www.opengeospatial.org/standards>. Abgerufen am 17.08.2015
- Sagl G, Resch B, Blaschke T (2015) Contextual Sensing: Integrating Contextual Information with Human and Technical Geo Sensor Information for Smart Cities. Sensors 2015 (15): 17013-17035
- SAP PAL (2015) SAP HANA Predictive Analysis Library. [http://help.sap.com/hana/sap\\_hana\\_predictive\\_analysis\\_library\\_pal\\_en.pdf](http://help.sap.com/hana/sap_hana_predictive_analysis_library_pal_en.pdf). Abgerufen am 17.08.2015
- Schachter EN et al. (2015) Outdoor air pollution and health effects in urban children with moderate to severe asthma. Air Quality, Atmosphere & Health 2015 (4):1-13
- Sow D et al. (2013) Mining of Sensor Data in Health Care: A Survey. In: Aggrawal CC (Hrsg) Managing and Mining Sensor Data. Springer, Heidelberg
- Spring XD (2015) Spring XD. <http://projects.spring.io/spring-xd>. Abgerufen am 17.8.2015
- UBA (2014) Umweltbundesamt, Hintergrund Luftqualität 2013. Bericht des Umweltbundesamtes, 2014
- 52n SOS (2015) Sensor Web Community. <http://52north.org/communities/sensorweb/index.html>. Abgerufen am 17.8.2015



# Matchmaking in der Sharing Economy mit Hilfe von semantischen Technologien

Tizian Böger<sup>1</sup> und Moritz von Hoffen<sup>2</sup>

<sup>1</sup> Westfälische Wilhelms-Universität Münster, Institut für Wirtschaftsinformatik,  
tizianboeger@uni-muenster.de

<sup>2</sup> Westfälische Wilhelms-Universität Münster, European Research Center for Information  
Systems (ERCIS), moritz.von.hoffen@ercis.uni-muenster.de

## Abstract

Die Sharing Economy ist ein viel diskutiertes Phänomen, bei dem sogenannte „Peers“ Ressourcen untereinander teilen. In den letzten Jahren sind in diesem Geschäftsfeld viele siloartige Plattformen entstanden. Dies erschwert es potentiellen Nutzern, passende Angebote zu finden. In diesem Artikel wird ein Ansatz vorgestellt, bei dem semantische Technologien für den Matchmaking-Prozess zum Einsatz kommen. Das kann dazu beitragen, anstatt vieler isolierter Plattformen, einen übergreifenden Marktplatz zu etablieren, bei dem die Beteiligten ihre Angebote und Präferenzen in Form von semantischen Konstrukten angeben. Hierzu wird eine entsprechende Architektur und eine zugehörige Ontologie konzipiert und deren prototypischer Einsatz demonstriert.

## 1 Einleitung

Die Sharing Economy ist ein noch recht junges und doch schon viel diskutiertes Phänomen, welches vor allem durch moderne Technologien ermöglicht wird (vgl. Malhotra und van Alstyne 2014; Botsman 2015). Ein Beispiel hierfür ist die Unterkunft-Vermittlungsplattform airbnb<sup>1</sup>. Hierbei können Privatpersonen Wohnraum von lokalen Gastgebern für einen bestimmten Zeitraum gegen Geld mieten. Damit macht airbnb nicht nur dem traditionellen Hotelgewerbe Konkurrenz, es stellt auch ein Problem für den regulären Wohnungsmarkt dar: Denn inzwischen werden vermehrt Wohnungen nur noch angemietet, um sie auf airbnb anzubieten und so einen maximalen Profit zu erwirtschaften (vgl. Henwood 2015).

In diesem Zusammenhang ist es besonders interessant, dass viele Unternehmen in der Sharing Economy nur noch als reine Matchmaking-Plattformen agieren, über die Angebot und Nachfrage zusammenführt werden (vgl. Linne et al. 2014, S. 9). Es ist für diese Unternehmen daher besonders wichtig, einen effizienten und präzisen Matchmaking-Prozess abzubilden. Dieser Vorgang ist Gegenstand der anschließenden Ausarbeitung. Hierzu wird ein Prototyp entwickelt, der eine

---

<sup>1</sup> <https://www.airbnb.de>

Vermittlungsplattform darstellt. Die Besonderheit liegt dabei in der Benutzung von semantischen Technologien zum Zusammenführen von Angebot und Nachfrage.

### 1.1 Problemstellung

Die Unternehmen in der Sharing Economy können als zweiseitige Märkte angesehen werden (vgl. Hu und Zhou 2015, S. 2). MALHOTRA UND VAN ALSTYNE (2014) sprechen hierbei von „internet mediaries“, also Online-Vermittlungsdiensten, die das Zusammenführen von Angebot und Nachfrage global und in Echtzeit ermöglichen. Dabei basiert das Matchmaking bei vielen Plattformen auf Faktoren wie beispielsweise dem derzeitigen Standort (z. B. Uber.com) oder dem übereinstimmenden Zeitraum und Anzahl der Betten (z. B. airbnb). Hierbei findet keine semantische Zusammenführung über beispielsweise implizit gegebene Präferenzen statt. So könnte etwa die Suche bei einer Unterkunftsvermittlung wie airbnb nach einer partytauglichen Wohnung in Berlin implizit zusätzlich bedeuten, dass die gesuchte Wohnung in einem Szeneviertel von Berlin liegen soll und darüber hinaus lautstärketolerante Nachbarn hat. Derartige Rückschlüsse lässt airbnb derzeit nicht automatisch und auch nur begrenzt manuell über Filter zu. Ein weiteres Problem stellt das Austreten aus dem Markt von „unmatched demand and supply“ dar (vgl. Hu und Zhou 2015, S. 3). Wenn beispielsweise ein potentieller Fahrgast über den Fahrdienstvermittler Uber keine passende Mitfahrgelegenheit findet, wechselt er zu einem klassischen Taxianbieter und geht als Kunde für Uber verloren.

Dieser Artikel versucht, die unterschiedlichen individuellen Präferenzen oder auch „taste[s]“ (vgl. Hu und Zhou 2015, S. 3) von Angebot und Nachfrage mithilfe von semantischen Technologien zusammenzubringen. Dabei liegt der Fokus insbesondere auf Gesuchen, die nicht direkt mit einem Angebot verknüpft werden können. Zu diesem Zweck orientiert sich dieser Artikel an den folgenden zwei Fragen:

- Wie können semantische Technologien den Matchmaking-Prozess unterstützen?
- Kann die Sharing Economy von diesem semantischen Matchmaking profitieren?

Zum Beantworten wird der Matchmaking-Prozess mithilfe von semantischen Technologien anhand des Beispiels eines Naturalienmarktplatzes demonstriert. Der hier entwickelte Prototyp enthält somit Naturalienangebote, wie z. B. Orangen von verschiedenen Bauern, die dann über eine Suchfunktion mit einem Nachfrager zusammengeführt werden.

## 2 Stand der Forschung

### 2.1 Semantic Web-Technologien

Das Semantic Web bringt Struktur in den Inhalt des Web und ermöglicht es, Daten anzureichern und mit anderen Informationen in Beziehung zu setzen (vgl. Berners-Lee et al. 2001). Bisher herrscht das Problem, dass Websites ihren Inhalt zusammen mit der Präsentation des Inhalts vermischen, was es für Maschinen schwierig macht, den Inhalt einer Seite zu verstehen. Genau dort setzt das Semantic Web an: Wenn Websites in strukturierter Form vorliegen, können Maschinen den Inhalt dieser verstehen. Damit wird es möglich, dass sie Inhalte automatisch verarbeiten und für andere Zwecke nutzen können. Darüber hinaus bietet das Semantic Web eine Infrastruktur, um unterschiedliche Wissensbasen und Datenbanken zu verknüpfen und hierdurch Wissen zu aggregieren und zu verknüpfen (vgl. Bizer et al. 2009). Es findet somit auch eine



Wissensverwaltung und ein Informationsaustausch statt, welcher vorrangig durch Ontologien ermöglicht wird.

Die auf der Web Ontology Language (OWL) basierende GoodRelations Ontologie dient dem semantischen Beschreiben von Produkten, Dienstleistungen und deren Angebote im Web (Hepp 2008). Damit bietet GoodRelations erstmals eine Ontologie und Infrastruktur „for [a] real Semantic Web-based e-commerce on a Web scale“ (vgl. Hepp 2008). Eine Vielzahl von Unternehmen und Online-Händler wie z. B. Google oder BestBuy.com nutzen diese Ontologie und verhelfen ihr so zum zweiten Platz der meist genutzten Ontologien im Web. Mit GoodRelations ist es darüber hinaus möglich, Handelsstrukturen und -daten über Klassen und Beziehungen darzustellen (vgl. Hepp 2008). Dazu zählen neben den oben genannten Produkten, Dienstleistungen und Angeboten beispielsweise auch die Beschreibung von Firmen und Geschäften sowie genaue Produktbeschreibungen über z. B. Mengenangaben (vgl. Hepp 2008).

## 2.2 Sharing Economy

Die Sharing Economy kann in einen produzierenden und in einen konsumierenden Bereich unterteilt werden. Während der produzierende vor allem die gemeinschaftliche Produktion von Wissen beinhaltet (z. B. Wikipedia, Open Source Software), werden im konsumierenden Bereich vorrangig Güter und Dienstleistungen mit anderen geteilt. Im Konsumbereich liegt somit der Fokus von Sharing Economy-Plattformen auf dem Unterstützen und Koordinieren von Transaktionen (vgl. Allen und Berg 2014, S. 16).

Der Begriff Sharing Economy ist ein Neologismus und ist seit dem im Jahr 2010 veröffentlichten Buch „What’s Mine Is Yours“ von BOTSMAN UND ROGERS weiter verbreitet. Er wird in der Literatur auch mit „Collaborative Consumption“ (vgl. Botsman und Rogers 2011), „Collaborative Economy“ (vgl. Owyang et al. 2013), „Peer Economy“ (vgl. Botsman 2013), „the Mesh“ oder „Product Service Systems“ (Botsman und Rogers 2011, S. 71 f.) umschrieben. Dabei umfasst jeder dieser Begriffe leicht unterschiedliche Bereiche, die allerdings viele Überschneidungspunkte besitzen (vgl. Botsman 2013).

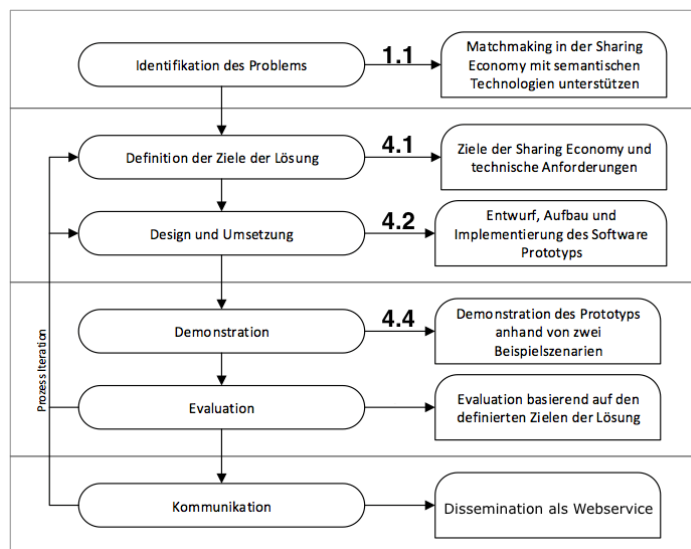
Der Begriff Collaborative Consumption (CC) wurde erstmals 1978 von FELSON UND SPAETH (1978) eingeführt. Sie beschreiben die CC als „those events in which one or more persons consume economic goods or services in the process of engaging in joint activities with one or more others“ (vgl. Felson und Spaeth 1978). BOTSMAN (2015) geht hingegen von einem ganz anderen Ansatz aus: Sie definiert den kollaborativen Konsum als eine durch Technologie gestützte „reinvention of traditional market behaviors“, unter anderem beim Mieten, Tauschen und Schenken. Somit unterscheidet sich ihre Definition von FELSON UND SPAETH (1978), da bei FELSON UND SPAETH (1978) als Beispiel auch „drinking beer with friends“ zur CC gezählt wird. Die von BELK (2014) vorgeschlagene Definition für den kollaborativen Konsum unterscheidet sich wiederum von den beiden vorherigen. Er nennt „coordinating the acquisition and distribution of a resource for a fee or other compensation“ zwischen Menschen als Aufgabe der CC.

Die Collaborative Economy (CE) wird von OWYANG ET AL. (2013, S. 4) definiert als „economic model where ownership and access are shared between corporations, startups, and people“. Sie argumentieren weiter, dass daraus Markteffizienzen entstehen, die „new products, services, and business growth“ hervorbringen. Dazu gehören die Bereiche „resell, rent, subscription, co-own, swap and gift“. BOTSMAN (2015) konkretisiert diese Definition für die CE, indem sie statt von einem Teilen von „ownership and access“, (vgl. Owyang et al. 2013, S. 4), nun von einem „matching [of] needs and haves [of underused assets]“ spricht. Darüber hinaus definiert BOTSMAN (2015),

dass die CE aus dezentralisierten Netzwerken besteht, in denen „traditional middlemen“ umgangen werden.

### 3 Methodik

Ziel dieses Artikels ist das Erschaffen eines Produkts in Form eines Softwareprototyps und der Beschreibung des dazugehörigen Entwicklungs- bzw. Designprozesses. Design Science unterstützt die Verwirklichung menschlicher Ziele und wird verwendet, um innovative Lösungen zu erschaffen (vgl. March und Smith 1995). Hierbei liegt der Fokus auf der Beseitigung ungelöster Probleme oder der effizienteren Gestaltung bereits gelöster Probleme (vgl. Hevner et al. 2004). Da in dieser Ausarbeitung ein „Artifact“ (vgl. Hevner et al. 2004; March und Smith 1995; Peffers et al. 2007) für „human purposes“ (vgl. March und Smith 1995) entwickelt wird, orientiert sich der Aufbau dieses Artikels an der Design Science Research-Methode (DSRM).



**Bild 3.1: Design Science-Prozess nach PEFFERS ET AL. (2007) mit Ausprägungen in dieser Arbeit**

Bild 3.1 fasst die verschiedenen Aktivitäten zusammen, die durchgeführt wurden, um das Artefakt zu konzipieren und implementieren. Während des Entwicklungsprozesses wurde der Prototyp durch eine parallel laufende Evaluation iterativ angepasst und um Funktionen erweitert. Im Rahmen dieses Artikels werden insbesondere der Entwurf und Aufbau der Lösung, aber auch die Demonstration selbiger, behandelt.

### 4 Systementwicklung

Dieses Kapitel liefert einen Überblick über den entwickelten Softwareprototyp. Die Struktur orientiert sich dabei an der im Kapitel 3 erläuterten Design Science Research-Methode.

Zunächst werden die Anforderungen an den Prototypen definiert, um dann auf Basis derer einen Entwurf der Software zu skizzieren. Darauf basierend findet im Anschluss die Realisierung des Prototyps statt. Hierbei wird als Erstes die Struktur der für den Prototyp aufgestellten Ontologien vorgestellt, um schließlich den eigentlichen Matchmaking-Prozess detailliert zu demonstrieren.

#### 4.1 Anforderungen und Zieldefinition

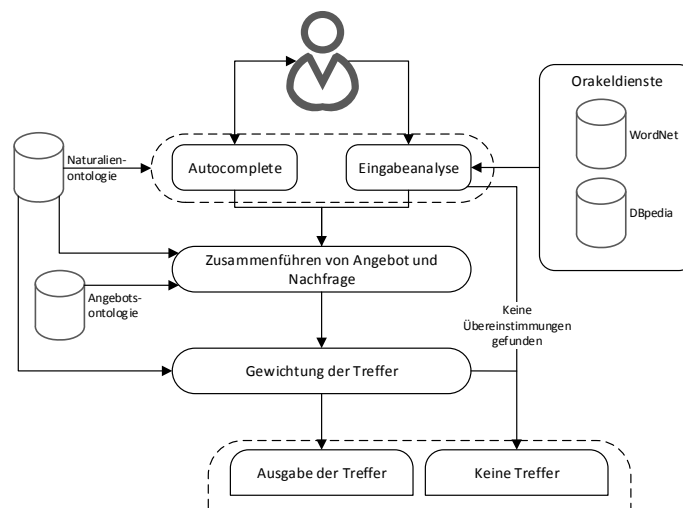
Die Entwicklung des Softwareprototyps orientiert sich an vorher definierten Zielen und Anforderungen. Er soll somit ein Matchmaking über semantische Beziehungen ermöglichen und damit einhergehend die individuellen Präferenzen der Teilnehmer aufnehmen und verarbeiten können.

Der Grundsatz des dezentralen Marktplatzes soll einen zentralen Aspekt des Prototyps darstellen. Angebote von überall auf der Welt sollen auf der entwickelten Plattform gebündelt angezeigt und mit individuellen Nachfragen zusammengeführt werden können.

#### 4.2 Entwurf

Der Entwurf des Softwareprototyps greift verschiedene in der Literatur von GIUNCHIGLIA ET AL. (2012) und PAREDES-VALVERDE ET AL. (2015) präsentierte Ansätze zum semantischen Matchmaking auf und verbindet diese mit selbst entwickelten Ideen. Weiterhin integriert er einige der in Kapitel 4.1 definierten Ziele und Anforderungen. Der Entwurf besteht aus insgesamt vier verschiedenen Modulen, sowie zwei Ontologien und zusätzlichen Diensten. Dabei dienen die ersten drei Module dem Matchmaking-Prozess, während das vierte die Ergebnisse des Prozesses visualisiert. Im Nachfolgenden wird der Entwurf basierend auf Bild 4.1 näher erläutert.

Die Naturalienontologie beinhaltet unterschiedliche Klassen an Naturalien wie z. B. Apfel oder Gurke und bildet somit den Grundstamm an Daten, welche nachgefragt und angeboten werden können. Auf ihrer Basis findet der Matchmaking-Prozess statt. Die zweite verwendete Ontologie beinhaltet Angebote von unterschiedlichen Personen und wird für das Zusammenführen von Angebot und Nachfrage benötigt.



**Bild 4.1: Entwurf des Softwareprototyps**

Das erste Modul im Matchmaking-Prozess beschreibt die Eingabeverarbeitung. Dieses Modul nimmt die Anfrage eines Nutzers entgegen und vergleicht sie mit den Klassen aus der Naturalienontologie. Hierbei kann der Nutzer entweder über eine Autocomplete-Funktion die ihm vorgeschlagenen Naturalien für das anschließende Zusammenführen von Angebot und Nachfrage auswählen oder seine Suche unabhängig von den Autocomplete-Vorschlägen mit seiner selbstformulierten Anfrage ausführen. Im zweiten Fall wird die Sucheingabe dann über die Orakeldienste, bestehend aus WordNet und DBpedia, angereichert und versucht, eine passende

Klasse in der Naturalienontologie zu finden. Dieser Vorgang orientiert sich an den „Oracle-Services“ aus dem von GIUNCHIGLIA ET AL. (2012) präsentierten S-Match-Ansatz. Hier werden ebenfalls Dienste wie WordNet verwendet, um die Nutzereingabe zu verstehen. Auch das „Question Processing“ aus dem von PAREDES-VALVERDE ET AL. (2015) vorgestellten Onli-Framework hat den Entwurf dieses Moduls beeinflusst, da hier ebenfalls zunächst versucht wird, die Nutzereingabe zu verstehen, um sie im Anschluss zu klassifizieren. Sollte die angereicherte Nutzeranfrage trotzdem mit keiner Klasse aus der Naturalienontologie zusammengeführt werden können, ist der Matchmaking-Prozess beendet und die nächsten zwei Module werden übersprungen.

Im zweiten Modul findet der eigentliche Matchmaking-Prozess statt. Hierbei wird die mit einer validen Naturalienklasse übereinstimmende Nutzereingabe aus dem vorherigen Modul (Nachfrage) mit den vorhandenen Angeboten zusammengeführt. Daraus resultiert eine Liste mit direkten Treffern, bei denen Eingangsprodukt und Angebotsprodukt syntaktisch oder semantisch übereinstimmen und indirekten Treffern, welche über gemeinsame Oberklassen aus der Naturalienontologie zusammengeführt werden konnten.

Der letzte Schritt im Matchmaking-Prozess beinhaltet das Einordnen der aus Modul zwei übermittelten Treffer. Hierbei werden die Treffer in dem Prototyp nach diversen Kriterien bewertet und eingeordnet. Zum Einordnen der Treffer werden die Matches hier ebenfalls u. a. anhand ihrer Eigenschaften bewertet. Diese werden allerdings nicht von Anbieter oder Nachfrager festgelegt bzw. in der Suche übergeben, sondern sind in der Naturalienontologie bereits jeder Klasse fest zugeschrieben. Darüber hinaus findet eine Einordnung basierend auf SKOS Beziehungen, Hyponymen, Gewichtsangaben, gemeinsamen Oberklassen und Matchdistanzen statt.

Das letzte Modul stellt die graphische Komponente des Prototyps dar. Hier werden die Ergebnisse aus dem Matchmaking-Prozess für den Nachfrager visualisiert. Zu diesem Zweck werden die nach Relevanz sortierten Ergebnisse absteigend dem Nutzer präsentiert. Zusätzlich zu der übereinstimmenden Naturalienklasse werden der Anbieter des Produkts sowie die erreichte Punktzahl, welche die Positionierung des Treffers rechtfertigt, angezeigt. Diese weiterführenden Informationen schaffen Transparenz auf der Plattform und sollen das Vertrauen der Nachfrager stärken. Bei keinen vorhandenen Treffern erhält der Nutzer eine entsprechende Nachricht.

### 4.3 Verwendete Ontologien

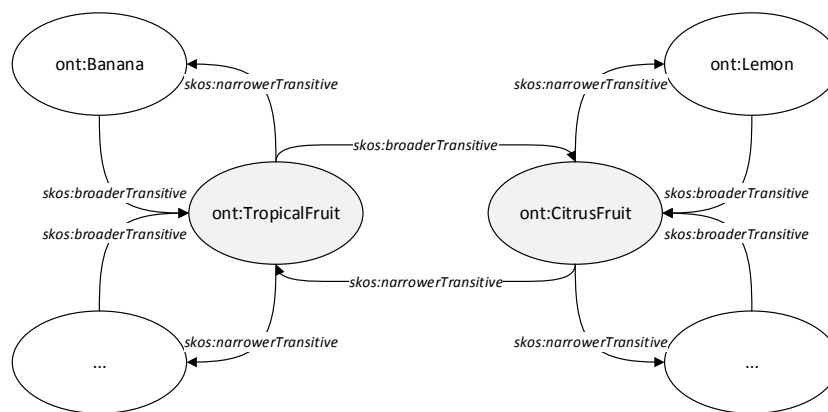
Die für den Prototyp verwendeten Ontologien basieren beide auf dem RDF/XML-Format. Sie sind auf einem externen Server gespeichert um einen leichten Zugriff über einen klar definierten Namespace zu ermöglichen.

Die Naturalienontologie nimmt im Prototyp eine wichtige Rolle ein. Sie wird für den gesamten Matchmaking-Prozess benötigt und umfasst die Menge aller Produkte, die im Prototyp angeboten oder nachgefragt werden können. Um die Produktvielfalt möglichst breit zu gestalten, bildet die Basis für die Naturalienontologie ein Ausschnitt aus den SUMO-Domänenontologien. Dazu wurden aus der in OWL übersetzten SUMO-Variante die Klassen „ReproductiveBody“ und „PlantAgriculturalProduct“ zusammen mit fast allen dazugehörigen Unterklassen herausgelöst. Die gewählten Oberklassen stehen für pflanzliche und landwirtschaftlich erzeugte Produkte und eignen sich daher gut zur Repräsentation von Naturalien.

Im nächsten Schritt findet eine Optimierung der Naturalienontologie im Hinblick auf den Fokus des Matchmaking statt. Hierzu wurden zunächst einige der SUMO-spezifischen Eigenschaften, wie z. B. *subsumingRelation*, aus den vorhandenen Klassen entfernt. Anschließend wurden die

Beziehungen zwischen den einzelnen Klassen bearbeitet. Dabei wurden insbesondere alle Klassen mit mehr als einer *rdfs:subClassOf*-Beziehung untersucht. Bei bestehenden transitiven Beziehungen zwischen den in einer Klasse vorhandenen Oberklassen wurde die jeweils höher gestellte entfernt. Beispielsweise besitzt die Klasse „Apple“ nun nur noch die Oberklasse „Fruit“ und verweist nicht mehr gleichzeitig auf die Klasse „FruitOrVegetable“, da die Zugehörigkeit zu dieser Klasse über „Fruit“ bereits transitiv gegeben ist.

Die Ontologie wurde weiterhin um einige Naturalien ergänzt, um die Vielfalt der möglichen Angebote zu erweitern. So wurden z. B. Erdbeeren und Himbeeren hinzugefügt. Für den Matchmaking-Prozess wurden zudem semantische Beziehungen mittels SKOS (*skos:semanticRelation*) ergänzt. Dazu wurden die Eigenschaften *skos:exactMatch*, *skos:narrowMatch*, *skos:broadMatch*, *skos:narrowerTransitive* und *skos:broaderTransitive* ausgewählten Naturalien beigelegt. Beispielsweise ist eine Zitrone („Lemon“) nun mit einer Limone („Lime“) über *skos:narrowMatch* verbunden. In der Ontologie dient die Beziehung zwischen den Oberklassen „TropicalFruit“ und „CitrusFruit“ als Beispiel für die Umsetzung von transitiven Relationen in diesem Prototyp. Hier stehen jeweils die zugeordneten Naturalien in einer *skos:narrowerTransitive* und *skos:broaderTransitive*-Beziehung mit ihrer Oberklasse.



**Bild 4.2:** Darstellung der transitiven Beziehungen in einem Ausschnitt der Naturalienontologie

Zusätzlich können Naturalien bis zu zwei verschiedene Geschmackseigenschaften (*tastes*) besitzen. Dazu zählen die Charakteristiken „sweet“, „sour“, „savory“, „bitter“ und „hot“. Der Zustand oder die *condition*, in der die Früchte verzehrt werden, kann entweder „fresh“ oder „dry“ sein. Ähnlich wie die SKOS Beziehungen spielen diese Eigenschaften eine wichtige Rolle beim Matchmaking.

Die zweite Ontologie des Prototyps beschreibt die verfügbaren Angebote. Hierbei dient das GoodRelations-Vokabular als Basis zur Beschreibung der Angebote. Ein Angebot besteht immer aus einer *gr:BusinessEntity*, in welcher der Name und Herkunftsort des Anbieters sowie mindestens ein Verweis auf ein *gr:Offering* (ein angebotenes Produkt) gespeichert werden. In einem Angebot (*gr:Offering*) ist sowohl der Geschäftszweck über *gr:BusinessFunction* abgebildet als auch eine nähere Beschreibung des angebotenen Produkts über den *gr:TypeAndQuantityNode* gegeben. Ein *gr:TypeAndQuantityNode* besteht dabei aus dem Typ des Produkts mit Verweis auf die Naturalienontologie (*gr:typeOfGood*), der Menge (*gr:amountOfGood*) und der Mengeneinheit (*gr:hasUnitOfMeasurement*) des Produkts.

Die Repräsentation des *gr:Offerings* wurde dabei bewusst einfach gehalten. So verfügt jedes Angebot nur über die für diesen Prototyp notwendigen Eigenschaften. Dazu verweist

*gr:typeofGood* im *gr:TypeAndQuantityNode* beispielsweise direkt auf eine Klasse der *Naturalienontologie*, anstatt beispielsweise auf eine Instanz des jeweiligen Produkts zu referenzieren.

#### 4.4 Gewichtung von Matches

Um ein Ranking unterschiedlicher Matches zu ermöglichen, wird ein einfaches Punktesystem verwendet, welches im Folgenden beschrieben wird.

Sofern in der Nachfrage keine Mengenangabe (z.B. in kg) definiert ist, erhalten alle direkten Matches im Ranking-Prozess die maximale Punktzahl, da hier Angebot und Nachfrage identisch sind. Bei einer Nachfrage mit angegebener Menge erhalten nur direkte Matches mit ausreichender Angebotsmenge die maximale Punktzahl. Treffer mit syntaktisch übereinstimmenden Ontologieklassen, aber einer zu niedrigen Angebotsmenge, werden dem Kriterienkatalog in Tabelle 4.1 entsprechend ein Punkt von der maximalen Punktzahl abgezogen.

Kriterien	Punktzahl
Naturalien spezifische Eigenschaften:	
- taste	+2 pro Geschmack
- grow	+1
- condition	+1
SKOS Beziehungen:	
- narrowMatch	+4
- broadMatch	+3
- narrowerTransitive	+2
- broaderTransitive	+1
Match-Distanz:	
- Gleiche direkte Oberklasse	+3
- Distanz kleiner als vier	+1
Mengenangabe vorhanden:	
- Nachgefragte Menge kleiner gleich angebotene	+1
- Nachgefragte Menge größer als angebotene	- 1
Übereinstimmendes Hyperonym	+1

**Tabelle 4.1: Kriterienkatalog zur Gewichtung der Treffer**

Alle übrigen Matches werden schrittweise anhand des Kriterienkatalogs aus Tabelle 4.1 bewertet. Als Vergleichsbasis dient dabei das nachgefragte Produkt. So werden zunächst Punkte für übereinstimmende naturalienspezifische Eigenschaften vergeben. Da der Geschmack eine wichtige Unterscheidungsfunktion bei der Einordnung von Naturalien bildet, erhält dieser einen Punkt mehr als die *grow*- und *condition*-Charakteristika. Wenn Nachfrageprodukt und Treffer also z.B. einen übereinstimmenden Geschmack haben, erhält der Treffer zwei Punkte. Ähnlich wie bei der Berechnung der maximalen Punktzahl gibt es bei der Gewichtung der Treffer auch logische Restriktionen: Für den Fall, dass der aktuell zu gewichtende Match über *skos:narrowMatch* oder *skos:broadMatch* mit der Nachfrage verbunden ist, wird auf die Überprüfung von *skos:narrowerTransitive* und *skos:broaderTransitive* Relationen verzichtet. Ebenfalls fällt in diesem Fall die Bewertung der Match-Distanz weg, da die möglichen zusätzlichen Punkte das Gesamtergebnis des Matches verzerren würden. Weiterhin findet eine Untersuchung auf eine Match-Distanz kleiner als vier nur dann statt, wenn weder *skos:narrowMatch* oder *skos:broadMatch* Beziehungen vorhanden sind, noch der aktuelle Match bereits über die direkte Oberklasse mit der Nachfrage verbunden ist. Andernfalls würde dieser zusätzliche Punkt die Positionierung des Matches im Vergleich zu den anderen Treffern unverhältnismäßig verbessern. Besteht dahingegen eine transitive Beziehung (*skos:narrowerTransitive/skos:broaderTransitive*)

zu einer anderen Oberklasse, wie beispielsweise zwischen tropischen Früchten und Zitrusfrüchten (siehe Bild 4.2), ohne gleichzeitige direkte SKOS-Relationen (*skos:narrowMatch* oder *skos:broadMatch*) und ohne bereits für Match-Distanzen vergebene Punkte, werden zu dem Treffer entsprechend ein bzw. zwei Punkte addiert. Da SKOS-Beziehungen hier die eindeutige Verwandtschaft von Naturalien beschreiben, werden im Falle einer Übereinstimmung die meisten Punkte vergeben.

Der letzte Punkt des Kriterienkatalogs behandelt Hyperonyme. Mithilfe von WordNet werden diese Oberbegriffe für den jeweils aktuellen Match und das Nachfrageprodukt einzeln ermittelt und auf mindestens eine Übereinstimmung hin untersucht. Beinhalten beide Produkte ein gleiches Hyperonym, erhöht sich die Punktzahl des Matches um eins.

#### 4.5 Demonstration

In diesem Abschnitt wird anhand von zwei Beispieleingaben die Funktionalität des Prototyps vorgestellt. Dabei liegt der Fokus vor allem auf der Gewichtung der Ergebnisse, da dies die zentrale Funktion zur Unterscheidung der Trefferqualität darstellt. Um die Ergebnisse des Matchmaking-Prozesses besser nachvollziehen zu können, ist es wichtig, einen Überblick über die angebotenen Produkte zu besitzen. Zu diesem Zweck werden alle verfügbaren Angebote in Tabelle 4.2 aufgeführt.

Anbieter	Produkt	Menge
Xing Li	Rice grain	2000 kg
	Black pepper	400 kg
Hermann Bauer	Apple	300 kg
Abimbola Attah	Banana	650 kg
	Cocoa bean	250 kg
Valentina Pedroso	Lemon	700 kg
	Orange fruit	230 kg
Giovanni Montinari	Olive	970 kg
	Tomato	533 kg
Olina Macek	Cucumber	1000 kg
	Radish	710 kg
	Potato tuber	1460 kg

**Tabelle 4.2:** Liste aller im Prototyp vorhandenen Angebote

In dem ersten Beispielszenario sucht der Nutzer nach 30 Pfund Früchten. Dazu gibt er in das Suchfeld „30 pound fru“ ein. Bereits ab dem dritten Buchstaben des Produktnamens erhält der Nachfrager die in Bild 4.3 gezeigten Autocomplete-Vorschläge. Dabei fällt auf, dass die Gewichtseingabe in Höhe von 30 Pfund bereits in Kilogramm umgewandelt wurde.

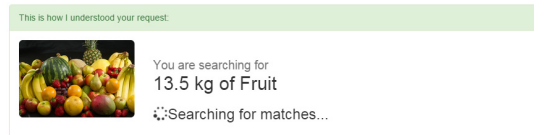
Enter your demand

30 pound fru Get offers!

- 13.5 kg of Fruit
- 13.5 kg of Pear Fruit
- 13.5 kg of Date Fruit
- 13.5 kg of Breadfruit
- 13.5 kg of Lemon Fruit

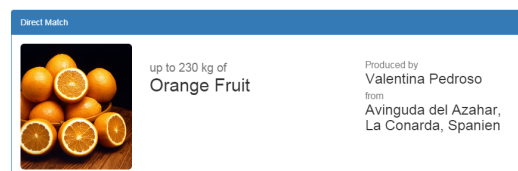
**Bild 4.3:** Ausschnitt der Autocomplete-Vorschläge für die Eingabe „fru“

Nachdem der Nutzer den Vorschlag für „13.5 kg of Fruit“ aus den Vorschlägen ausgewählt hat, erscheint ein Ladevorgang mit der Information, dass die Produktdaten zusammengestellt werden. Sobald das Nachfrageprodukt aus dem Backend eingetroffen ist, erscheint die Suchanfrage des Nutzers in einem grünen Panel mit einem Bild der entsprechenden Früchte (siehe Bild 4.4).



**Bild 4.4: Visualisierung des Nachfrageprodukts „Fruit“**

Zeitgleich beginnt die Suche nach Treffern im Backend. Nachdem die Ergebnisse aus dem Backend übermittelt wurden, wird die Anzahl der gefundenen Matches angezeigt. Ergebnisse werden unterhalb der Suchanfrage in einer Liste angezeigt. Neben dem eigentlichen Produkt werden auch Informationen zum Anbieter angezeigt (Bild 4.5). Bei der Suche nach Früchten gibt es nur direkte Treffer. Darunter fallen alle angebotenen Produkte, die zur Oberklasse Früchte gehören.



**Bild 4.5: Direkter Treffer in der Ergebnisliste**

Das nächste Beispiel behandelt die Suche nach Zucchini. Der Nutzer tippt hierfür „200 kg of courgette“ in das Suchfeld ein und bestätigt seine Suche mit einem Klick auf „Get offers!“. Der Begriff „courgette“ wird vor allem im Raum Großbritannien verwendet und ist nicht in der Naturalienontologie als Klasse enthalten. Aus diesem Grund erhält der Nutzer auch keine Autocomplete-Vorschläge. Nachdem die Anfrage über den Klick auf den Button abgeschickt wurde, erscheint eine Ladeanzeige mit der Information „Trying to understand your request...“. Im Backend wird zeitgleich eine passende Ontologiekategorie gesucht. Da es keine direkte Übereinstimmung gibt, sucht das System nach Synonymen und findet bei WordNet den Begriff Zucchini, welcher einer Ontologiekategorie entspricht.

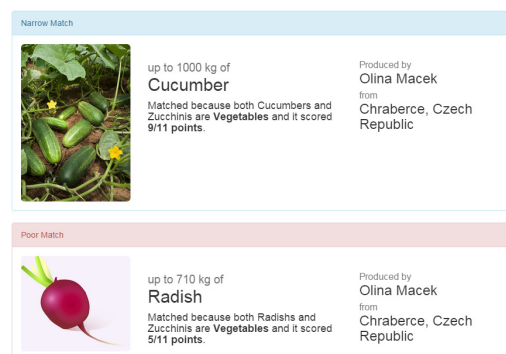
Im Anschluss wird das Nachfrageprodukt zu 200 kg Zucchini erstellt. Die maximale Punktzahl setzt sich hierbei aus der folgenden Bewertung zusammen: Zunächst erhält die Zucchini vier Punkte für je eine vorhandene Eigenschaft *grow* (ein Punkt), *taste* (zwei Punkte) und *condition* (ein Punkt). Weitere vier Punkte kommen durch die Möglichkeit eines *skos:narrowMatch* mit einer Gurke dazu. Da mit 200 kg eine Mengenangabe gesetzt wurde, erhält das Nachfrageprodukt einen weiteren Punkt. Über WordNet sind weiterhin Hyperonyme für Zucchini vorhanden. Daher erhöht sich die maximale Punktzahl um einen zusätzlichen Punkt. Abschließend kommt noch ein extra Punkt für direkte Matches dazu, sodass insgesamt elf Punkte von Treffern maximal erreicht werden können.

Das erstellte Nachfrageprodukt wird nun an das Fronend zurückgeschickt und dort, wie im vorhergegangenen Beispiel beschrieben, präsentiert. Für Zucchini gibt es keine direkten Matches, dafür aber sechs indirekte. In Bild 4.6 ist ein Auszug aus der Ergebnisliste zu sehen.



Die Gurke kann über eine *skos:narrowMatch*-Beziehung bereits vier Punkte sammeln und erhält zusätzlich durch den gleichen Geschmack (herzhaft), die gleiche Wachstumsart (auf dem Boden) und die gleiche Verzehrkondition (frisch) weitere vier Punkte.

Da sie weiterhin in einer ausreichenden Menge vorhanden ist, erreicht die Gurke schließlich neun von elf Punkten. Dem Radieschen fehlt hingegen die *skos:narrowMatch*-Beziehung und auch die Art des Wachstums unterscheidet sich von der der Zucchini. Dennoch erreicht es fünf Punkte, weil es über nur drei Kanten mit der Zucchini zusammengeführt werden kann und alle restlichen Eigenschaften übereinstimmen.



**Bild 4.6: Indirekte Treffer in der Ergebnisliste**

Die meisten verfügbaren Früchte, die im ersten Beispiel als direkte Treffer zurückgegeben wurden, fehlen in der Ergebnisliste für diese Suchanfrage. Sie besitzen zu wenige Übereinstimmungen mit der Zucchini und sind aus diesem Grund keine sinnvollen Treffer.

## 5 Diskussion

In dieser Arbeit wurde ein Software-Prototyp entwickelt, der die Möglichkeiten von semantischen Technologien im Matchmaking-Prozess aufzeigt. Dabei diente das Zusammenführen von Angebot und Nachfrage in der Sharing Economy als Grundmotivation, da in diesem Bereich viele Unternehmen als reine Matchmaking-Plattform operieren (vgl. Linne et al. 2014, S. 9). Es wurde vorgestellt, wie die Verwendung von Ontologien mit spezifischen Eigenschaften und SKOS-Relationen in Kombination mit einer Gewichtungsfunktion dazu beitragen können, die Trefferqualität zwischen Angebot und Nachfrage zu erhöhen.

Der entwickelte Prototyp stellt dabei eine sinnvolle Grundlage dar, welche die Vorteile der Verwendung semantischer Technologien im Matchmaking aufzeigt. Ein weiterer Ausbau der Funktionen und vor allem auch einer Kombination dieser mit bereits in der Sharing Economy eingesetzten Mitteln, kann einen noch effizienteren und präziseren Matchmaking-Prozess in Zukunft ermöglichen. Damit trägt diese Ausarbeitung der Forschung im Bereich des Matchmaking zwischen Angebot und Nachfrage in der Sharing Economy bei, indem sie neue Möglichkeiten für diesen Prozess aufzeigt und zusätzlich prototypisch umsetzt.

## 6 Literatur

Allen D, Berg C (2014) The sharing economy: How over-regulation could destroy an economic revolution. Institute of Public Affairs Australia. Australia. Online verfügbar unter

- [https://ipa.org.au/portal/uploads/Sharing\\_Economy\\_December\\_2014.pdf](https://ipa.org.au/portal/uploads/Sharing_Economy_December_2014.pdf), zuletzt geprüft am 14.09.2015.
- Belk R (2014) You are what you can access. Sharing and collaborative consumption online. In: *Journal of Business Research* 67 (8), S. 1595–1600. DOI: 10.1016/j.jbusres.2013.10.001.
- Berners-Lee T, Hendler J, Lassila, O (2001) The Semantic Web. In: *Scientific American* 284 (5), S. 34–43. Online verfügbar unter <http://www.sciam.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21>.
- Bizer C, Heath T, Berners-Lee T (2009) Linked data-the story so far. In: *Semantic Services, Interoperability and Web Applications: Emerging Concepts*, S. 205–227.
- Botsman R (2013) The Sharing Economy Lacks A Shared Definition. Online verfügbar unter <http://www.fastcoexist.com/3022028/the-sharing-economy-lacks-a-shared-definition>, zuletzt aktualisiert am 21.11.2013, zuletzt geprüft am 14.09.2015.
- Botsman R (2015) Defining The Sharing Economy: What Is Collaborative Consumption—And What Isn't? Online verfügbar unter <http://www.fastcoexist.com/3046119/defining-the-sharing-economy-what-is-collaborative-consumption-and-what-isnt>, zuletzt aktualisiert am 27.05.2015, zuletzt geprüft am 14.09.2015.
- Botsman R, Rogers R (2011) *What's mine is yours. How collaborative consumption is changing the way we live.* Rev. and updated ed. London: Collins (Collaborative consumption).
- Felson M, Spaeth JL (1978) Community Structure and Collaborative Consumption: A Routine Activity Approach. In: *American Behavioral Scientist* 21 (4), S. 614–624. DOI: 10.1177/000276427802100411.
- Giunchiglia F, Autayeu A, Pane J (2012) S-Match: An open source framework for matching lightweight ontologies. In: *Semantic Web* 3 (3), S. 307–317. Online verfügbar unter <http://dblp.uni-trier.de/db/journals/semweb/semweb3.html#GiunchigliaAP12>.
- Henwood D (2015) What The "Sharing Economy" Takes. (Cover story). In: *Nation* 300 (7), S. 12–15. Online verfügbar unter <http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=100798672&lang=de&site=ehost-live>.
- Hepp M (2008) GoodRelations: An Ontology for Describing Products and Services Offers on the Web. In: Aldo Gangemi und Jérôme Euzenat (Hg.): *Knowledge Engineering: Practice and Patterns*, Bd. 5268. Berlin, Heidelberg: Springer Berlin Heidelberg (Lecture Notes in Computer Science), S. 329–346.
- Hevner AR, March ST, Park J, Ram S (2004) Design Science In Information Systems Research. In: *MIS Quarterly* 28 (1), S. 75–105. Online verfügbar unter <http://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=12581935&lang=de&site=ehost-live>.
- Hu M, Zhou Y (2015) Dynamic Matching in a Two-Sided Market. In: *SSRN Journal*. DOI: 10.2139/ssrn.2592622.
- Keen A, Botsman R (2015) Can the Sharing Economy Provide Good Jobs? Some see benefits for many kinds of workers, others see a raw deal for most. Online verfügbar unter <http://www.wsj.com/articles/can-the-sharing-economy-provide-good-jobs-1431288393>, zuletzt aktualisiert am 10.05.2015, zuletzt geprüft am 01.09.2015.
- Linne M, Vogel HG, Dreyer A, Stampfl NS, Röder P (2014) *Smart Tourism - Share Economy im Tourismus. Produkte - Grenzen - Folgen.* Elmshorn: ITD-Verl. (Schriftenreihe Dienstleistungsmanagement, 15).
- Malhotra A, van Alstyne M (2014) The dark side of the sharing economy ... and how to lighten it. In: *Commun. ACM* 57 (11), S. 24–27. DOI: 10.1145/2668893.
- March ST, Smith GF (1995) Design and natural science research on information technology. In: *Decision Support Systems* 15 (4), S. 251–266. DOI: 10.1016/0167-9236(94)00041-2.
- Owyang J, Tran C, Silva C (2013) The Collaborative Economy. Products, services, and market relationships have changed as sharing startups impact business models. To avoid disruption, companies must adopt the Collaborative Economy Value Chain. Altimeter Group. USA.
- Paredes-Valverde MA, Rodríguez-García MÁ, Ruiz-Martínez A, Valencia-García R, Alor-Hernández G (2015) ONLI: An ontology-based system for querying DBpedia using natural language paradigm. In: *Expert Systems with Applications* 42 (12), S. 5163–5176. DOI: 10.1016/j.eswa.2015.02.034.
- Peppers K, Tuunanen T, Rothenberger MA, Chatterjee S (2007) A Design Science Research Methodology for Information Systems Research. In: *Journal of Management Information Systems* 24 (3), S. 45–77. DOI: 10.2753/MIS0742-1222240302.

# Improving the Quality of Multiple-Choice Exams by Providing Feedback from Item Analysis

Helena Bukvova<sup>1</sup>, Kathrin Figl<sup>1</sup>, and Gustaf Neumann<sup>1</sup>

<sup>1</sup> Vienna University of Economics and Business, {helena.lovasz-bukvova, kathrin.figl, gustaf.neumann}@wu.ac.at

## Abstract

Many universities use computer-assisted multiple-choice exams to handle large numbers of students in introductory courses. This paper describes how test-theoretical analysis can be used to offer continual feedback to examiners to ensure quality of multiple-choice exams. We present the design of an online system that gives examiners individual feedback on the test-theoretical quality of their multiple-choice exams in textual, tabular and graphical form, guiding them in item-revision and re-composition of exams if necessary.

## 1 Introduction

Many higher-education institutions face the challenge of assessing large numbers of students with only limited personal resources for scoring exams. One option to increase the efficiency of assessment is the use of standardised, structured exams that can be graded automatically. Among the most popular standardised assessment tools are multiple-choice (MC) exams. MC exams are a form of standardised, closed-question assessment that requires examinees to choose one or more correct answers from a collection of possible options. The type of MC question influences both their application in an exam as well as their grading (Haladyna 1992; Litzemberger et al. 2007). In its simplest form, an MC exercise consists of an item stem (e.g. a question or a problem description) and a set of answers presented to the examinees, instructing them to choose one answer that is true, whereas all others are false (distractors); this format is sometimes called single-choice question. It is also possible to include several correct options in the answer set, either informing the examinees about the number of correct options or allowing them to mark any number of options as correct (or even allow a none-of-the-above option). A slightly different format of MC questions also presents the examinees with several correct answers, but requires them to choose only one option that fits best. Further formats may require the examinees to indicate their level of confidence regarding the answer (Kolbitsch et al. 2008) or provide a brief explanation of their choice.

Advantages as the high level of standardisation and thus simplified (even automatic) grading (Chang et al. 2007) have led to a widespread use of MC-based assessment. Furthermore, the limited set of options allows MC questions to be answered comparatively quickly, thus allowing the examiners to assess a large number of areas within a limited time, providing high reliability of the

results (Sim and Rasiah 2006). Some studies also suggest that MC exams can be used to test higher-level knowledge (Haladyna 1992; Kastner and Stangl 2011). On the other hand, MC exams have been criticised with regard to their validity and fairness (McCoubrie 2004); more specifically, the evaluation of partial knowledge (Chang et al. 2007) and guessing are problematic issues. To address such shortcoming, prior research has developed approaches to identify or discourage guessing (Bereby-Meyer et al. 2002), mainly through grading algorithms (Bush 1999; Lesage et al. 2013).

While exam structure and grading approaches do influence the performance of MC exams, the quality of an MC exam is predominantly determined by the wording of the items (Cheung and Bucat 2002). The importance of suitably worded MC items leads to the fact that the quality of an MC exam strongly depends on the examiners' MC-writing abilities. Although examiners can be guided and supported by guidelines, writing MC items still remains a complex and creative activity (Rodriguez 2005). In order to improve their MC-writing skills, examiners need continuous feedback on the actual performance of their MC items (Möltner et al. 2006; Sim and Rasiah 2006). Such feedback would also be valuable to higher-education institutions that reuse the pools of MC items (McCoubrie 2004; Litzenger et al. 2007), as it would help them to identify MC items, which are suitable for further reuse or which should further be improved (McCoubrie 2004; Dickinson 2013).

To provide direct feedback on the quality of their MC items, we suggest that suitable performance indicators should be automatically calculated and made available to examiners via a learning management system. This paper describes our first insights from an ongoing project seeking to design such integrated examiner-oriented MC-feedback at the Vienna University of Economics and Business (WU). The research project can be characterized as design-oriented. Building on existing test-theoretical indicators, the main aim of the project is to create a supportive information system for examiners. In the following, we first discuss test-theoretical indicators that are useful to assess the quality of MC items and provide examiners with actionable feedback (see section 2). We then suggest a system that implements such feedback mechanisms directly into a learning management system (LMS) and demonstrate the use of the indicators with sample items (see section 3). Finally, in section 4, we discuss opportunities for further research in this context.

## 2 Test-theoretical analysis of MC indicators

In this section, we discuss the needs of MC examiners and give an overview on different facets of quality of MC items. We have considered the following criteria to describe a suitable MC item: (1) the item stem and answers must be understood by the examiner and the examinee in the closest possible manner, (2) the item must reflect the set, overall level of difficulty of the course, and (3) the item must display a sufficient level of challenge to allow the differentiation between different levels of knowledge among the examinees. There are two well-known and established standard indicators from classic test theory, which allow for an uncomplicated analysis of the item performance and provide examiners with clear, actionable feedback: item difficulty and discrimination index.

The *item difficulty* gives the percentage of students who have answered the item correctly (Bortz and Döring 2006). Thus, the term “item difficulty” might be confusing at first, because the higher the ex-post calculated indicator (the percentage of correct answers) is, the easier the item was. The item difficulty provides the examiners with a feedback on the performance of students on each item. Furthermore, this indicator helps to identify potentially falsely coded items, or topics insufficiently covered in the learning materials. If an MC exercise has more than one answer that is true, the

“correctness” of a MC task is not dichotomous. Therefore, besides providing the percentage of students who have answered the entire question (all answers) correctly, it is also useful to indicate the difficulty of each answer option. There are no exact boundaries for optimal item difficulty. Generally, items between 35% and 80% of difficulty are considered suitable (Sim and Rasiah 2006; Venter 2010).

The *discrimination index* is based on a division of the examinees into three groups: strong (top 27% examinees), weak (bottom 27% of examinees), and middle (Möltner et al. 2006). The discrimination index describes the difference of the relative number of correct answers of the strong and the weak group. The index helps examiners to identify items that are ambiguously worded or guessable, hence not suitable to differentiate between high and low performing examinees. Again, because students can have partially correct answers, it is also useful to provide the discrimination index for each single answer option as well as for an item overall. Test-theoretical literature suggests that the discrimination index should be above 0.2; items with a negative discrimination index lead to a lower validity of the test.

More indicators have been described in the literature, but we consider those indicators presented above to be the most effective ones in providing actionable feedback to examiners.

### **3 Design proposal for including MC item analysis in a learning management system**

While test-theoretical analysis for multiple-choice items is well established from a theoretical point of view, there are few learning management systems which support its use for instructors. Although test-theoretical analysis is used in research or to evaluate the performance of key exams such as PISA (e.g. Litzberger et al. 2007; Venter 2010), it is rarely available in day-to-day work of examiners (with exceptions, such as Wentzel 2006). Unlike constructed-response exams, MC exams are mostly used for large student groups in which performance problems of items (such as unclear wording or unsuitable difficulty) cannot be easily adjusted. Piloting is not an option for many MC exams due to time constraints and confidentiality issues. Frequent feedback and reflection are therefore crucial for quick improvement of MC-writing skills of examiners.

#### **3.1 Technical implementation**

Institutions that use MC exams to regularly assess large groups of students typically store and manage the MC items as well as the results (students' performance on the items in actual exams) digitally. This makes it possible to use the results to calculate item performance indicators. Figure 1 presents a diagram to illustrate how a learning management system can support the use of MC items in a teaching and learning setting based on a pool of MC items. Examiners add items to an MC-item pool and use the pool to generate new MC exams. Parts of the item pool can be opened to students to practise and prepare for exams. Students' exam results are connected to the MC items in the pool to calculate test-theoretical indicators for the MC items. The test-theoretical results are available for examiners (and can be made available for students) through a learning management system. The item pool is then updated with new information on the quality of items, providing feedback to the examiners and enhancing the quality of subsequent test, in which items are reused.

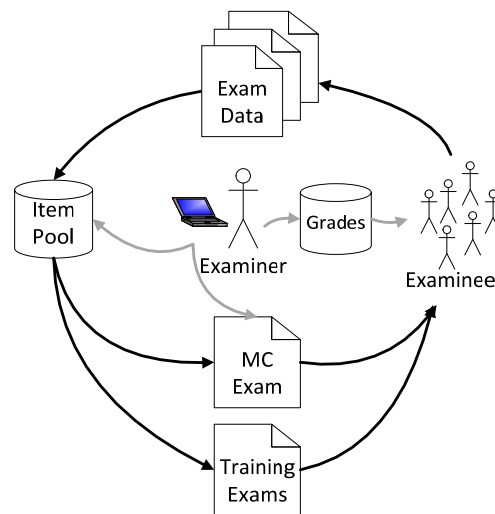


Figure 1. Learning management system to support MC exams

### 3.2 Exemplary application

In the following, we demonstrate the use of the indicators for exemplary items. The example items are taken from an MC exam assessing an introductory course of Business Information Systems at the Vienna University of Economics and Business. The exam contained 27 items and was taken by 451 students. Each MC item contained five answer options out of which between one and five were true (correct options are marked in the following tables with (+), distractors with (-)). The students were rewarded for marking correct answer options, and penalised for distractors. Thus, partial knowledge was rewarding and guessing discouraged. The MC exams were carried out in a written format and then scanned to allow for automated grading.

Performance group	Answer options					Overall
	a (+)	b (-)	c (-)	d (+)	e (-)	
Strong examinees (27%)	0.93	1.00	0.96	1.00	1.00	0.98
Middle examinees (46%)	0.84	0.99	0.78	0.94	0.99	0.91
Weak examinees (27%)	0.58	0.88	0.58	0.70	0.81	0.71
Item difficulty	0.79	0.96	0.77	0.89	0.95	0.87
Discrimination index	0.36	0.12	0.39	0.30	0.19	0.27

Table 1: Example item A

Table 1 shows the item difficulties of one example item for the three performance groups at the overall item and answer option level as well as the discrimination index (last line). The item difficulty indicators suggest that the item contains options that are too easy and thus, differentiate insufficiently between examinees. These are, in particular, the answer options *b* and *e*. These two distractors are identified as wrong by nearly all examinees. There are several possible explanations for such a result. Judging from the content of the specific item, it is likely that distractors could have been identified as wrong by most students based on common sense. The examiner would do well to replace or adapt the distractors. On the other hand, when correct options reach a high value for the item difficulty indicator, another possible explanation might be the fact that the learning goals concerning the item were achieved by all students.

Performance group	Answer options					Overall
	a (+)	b (-)	c (-)	d (+)	e (+)	
Strong examinees (27%)	0.38	0.51	0.33	0.81	0.71	0.55
Middle examinees (46%)	0.45	0.41	0.27	0.61	0.56	0.46
Weak examinees (27%)	0.56	0.53	0.23	0.54	0.51	0.47
Item difficulty	0.46	0.47	0.27	0.64	0.58	0.48
Discrimination index	-0.18	-0.02	0.10	0.27	0.20	0.08

Table 2: Example item B

Viewing the indicators in Table 2, it appears that two answer options (a, b) in this item can be solved by weak examinees better than by students with higher overall domain knowledge. This is not observable from the item difficulties, which mostly fits the 35%-80% range, but from the discrimination indices. The discrimination indices are too low for most answer options as well as overall; it is even negative for answer options *a* and *b*. The fact that weak examinees performed better than strong examinees on these two answer options can for instance indicate excessive guessing behaviour, or can be due to incorrect coding of the exercise. Since prior research has shown that weak examinees tend to guess more often than strong examinees (Sim and Rasiah 2006), this group might have an advantage with very difficult or unclear answer options. Another interpretation could be that very well prepared students realized potentially contradictory nuances of an item wording, while less prepared students just chose the answer that sounded right without reflecting it in detail. In this particular case, the item required a very detailed knowledge of encrypting methods; the examiners should consider revising the wording of the item. In general, the test-theoretical analysis cannot determine precise causes of undesired results, but the indicators provide hints to the examiner, which answer alternatives should be reconsidered. In some cases, it may also be appropriate to reflect and change the overall teaching approach for subject areas with low item difficulties or negative discrimination indices.

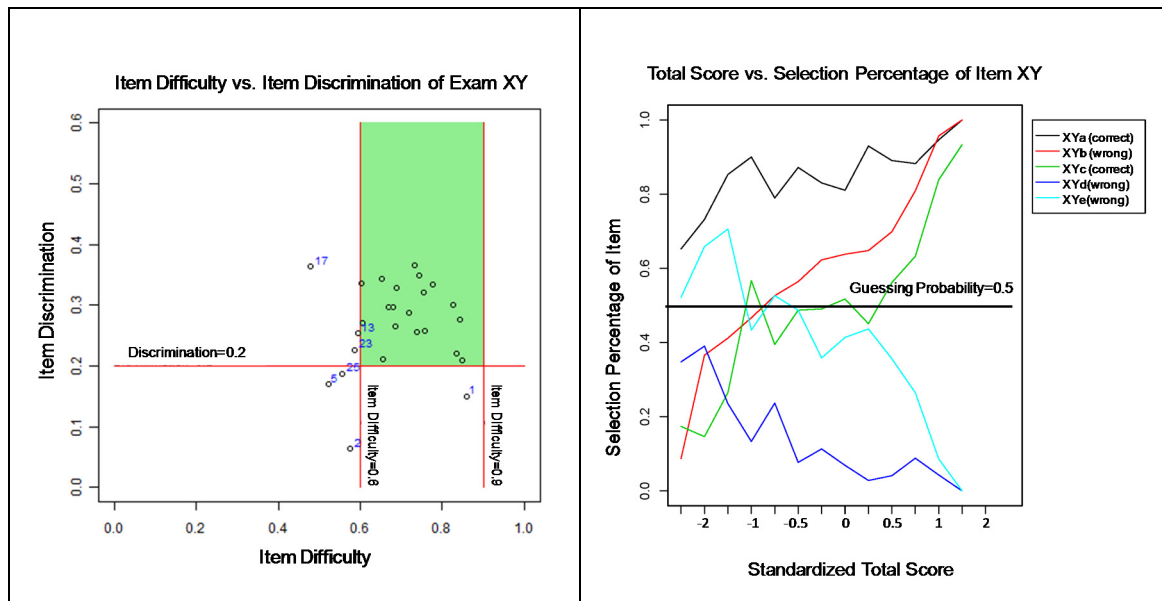


Figure 2: Item analysis graphics

Figure 2 shows two diagrams that visualize test-theoretical indicators. The graphic on the left highlights those items of an exam that fulfil test-theoretical quality criteria (in the coloured background) and those that have potential problems (white background). This diagram might be interesting to give an overview on item quality. The diagram on the right shows whether examinees choose to mark an item as correct depending on their overall performance in the MC test (standardized total score). Thus, the slope of the correct items should go from the lower left to the upper right (e.g. XYb), indicating that more higher performing than lower performing students solve it correctly, while the slope of wrong items should go from the upper left to the lower right of the graphic (e.g. XYe). Item XYa for instance is quite easy for all students, as the slope is not very steep. In addition, graphics on grade distribution and tests of normality of grade distribution could be provided by the system to evaluate whether a test differentiates well between students on all performance levels.

## 4 Conclusions and further research

The present project was undertaken to design an integrated system that automatically calculates a set of test-theoretical indicators for MC items based on exam results, providing examiners with a feedback on the quality of their MC items. The suggested system design is based on ongoing research efforts to support the process of studying for MC exams within the custom-developed learning management system Learn@WU (Andergassen et al. 2015). At the moment, the system provides the examiners with feedback on item difficulty and discrimination as well as on further details such as multidimensionality, item clustering, or item sensitivity. The provided indicators are not absolute measures per se, they do not brand an item as “good” or “bad” and they do not serve a self-purpose. Even items with indicators out of range can be correctly worded, querying essential knowledge from the course. It is certainly questionable, whether or not it is the purpose of a course and an exam to deliver good discrimination rather than knowledge. However, from our experience the indicators help to direct the examiners’ attention to problematic items in an exam, thus helping them reflect and improve both their MC-writing skill as well as their teaching (Talebi et al. 2013). Several possible directions for future research would be of interest. Opportunities exist for fellow scholars to conduct longitudinal research of item usage as well as to investigate the effect of online-learning behaviour (compare Andergassen et al 2014) on item performance.

## 5 Bibliography

- Andergassen M, Ernst G, Guerra V, Mödritscher F, Moser M, Neumann G, Renner T (2015) The Evolution of E-Learning Platforms from Content to Activity Based Learning. The Case of Learn@WU, In: International Conference on Interactive Collaborative Learning (ICL). Florence, Italy.
- Andergassen M, Mödritscher F, Neumann G (2014) Practice and Repetition during Exam Preparation in Blended Learning Courses: Correlations with Learning Results. *J Learn Anal* 1:48–74.
- Bereby-Meyer Y, Meyer J, Flascher OM (2002) Prospect Theory Analysis of Guessing in Multiple Choice Tests. *J Behav Decis Mak* 15:313–327. doi: 10.1002/bdm.417
- Bortz J, Döring N (2006) *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. Springer, Heidelberg, Germany.



- Bush M (1999) Alternative marking schemes for on-line multiple choice tests. 7th Annu Conf Teach 3–5.
- Chang SH, Lin PC, Lin ZC (2007) Measures of partial knowledge and unexpected responses in multiple-choice tests. *Educ Technol Soc* 10:95–109.
- Cheung D, Bucat R (2002) How can we construct good multiple-choice items? In: Science and Technology Educaiton Conference. Hong Kong.
- Dickinson JR (2013) How Many Options do Multiple - Choice Questions Really Have ? 171–175.
- Haladyna TM (1992) The Effectiveness of Several Multiple-Choice Formats. *Appl. Meas. Educ.* 5:73–88.
- Kastner M, Stangl B (2011) Multiple choice and constructed response tests: Do test format and scoring matter? *Procedia - Soc Behav Sci* 12:263–273. doi: 10.1016/j.sbspro.2011.02.035
- Kolbitsch J, Ebner M, Nagler W, Scerbakov N (2008) Can Confidence Assessment Enhance Traditional Multiple-Choice Testing? In: Interactive Computer Aided Learning, ICL2008. Villach, Austria.
- Lesage E, Valcke M, Sabbe E (2013) Scoring methods for multiple choice assessment in higher education - Is it still a matter of number right scoring or negative marking? *Stud Educ Eval* 39:188–193. doi: 10.1016/j.stueduc.2013.07.001
- Litzenberger M, Punter JF, Gnams T, et al (2007) Qualitätssicherung bei der Studierendenauswahl mittels lernpsychologisch fundierter Wissensprüfung. *Qualitätssicherung und -entwicklung an Hochschulen Methoden und Ergebnisse* 23–34.
- McCoubrie P (2004) Improving the fairness of multiple-choice questions: a literature review. *Med Teach* 26:709–712. doi: 10.1080/01421590400013495
- Möltner A, Schellberg D, Jünger J (2006) Grundlegende quantitative Analysen medizinischer Prüfungen. Basic quantitative analyses of medical examinations. *GMS Zeitschrift für Medizinische Ausbildung* 23:Doc53.
- Rodriguez MC (2005) Three options are optimal for multiple-choice items: A meta-analysis of 80 years of research. *Educ Meas Issues Pract* 24:3–13. doi: 10.1111/j.1745-3992.2005.00006.x
- Sim SM, Rasiah RI (2006) Relationship between item difficulty and discrimination indices in true/false-type multiple choice questions of a para-clinical multidisciplinary paper. *Ann Acad Med Singapore* 35:67–71.
- Talebi GA, Ghaffari R, Eskandarzadeh E, Oskouei AE (2013) Item Analysis an Effective Tool for Assessing Exam Quality, Designing Appropriate Exam and Determining Weakness in Teaching. 2:69–72. doi: 10.5681/rdme.2013.016
- Venter I (2010) Development of a Valid and Reliable Test for Higher-Educated Young Adults Measuring Dietary Fibre Food Source and Health-Disease Association Knowledge. *J Fam Ecol Consum Sci /Tydskrif vir Gesinsekologie en Verbruikerswetenskappe* 34:10–19. doi: 10.4314/jfec.v34i1.52906
- Wentzel C (2006) A review of INTEGRITY software: An online application to analyze multiple-choice tests and detect test-taking deception. *J Sci Educ Technol* 15:314–319. doi: 10.1007/s10956-006-9018-2



# **“IT via ERP”: A Novel Approach to Teaching IT Fundamentals to Business Administration students by Means of an Open Source ERP System**

**Mariela Castro Kohler<sup>1</sup> and Tobias Hagen<sup>1</sup>**

<sup>1</sup> Hochschule Offenburg, Fakultät Betriebswirtschaft und Wirtschaftsingenieurwesen,  
{mariela.castro, tobias.hagen}@hs-offenburg.de

## **Abstract**

First year Business Administration students tend to regard themselves as “non-computer scientists” and often have a lack of motivation about taking IT courses in general, either because they perceive them as too technical, too difficult or somewhat irrelevant. In an attempt to counteract this perception and increase the levels of engagement and willing attendance to class, we decided to flip the traditional lecture model and develop a new teaching and learning approach for the *IT Fundamentals* course using an open source Enterprise Resource Planning (ERP) system as the platform from which to draw the various underlying IT concepts and through which the relevant competences can be acquired.

This paper describes the implementation process of this new contextualized learning framework “IT via ERP” and the changes in the didactical methods to support it.

## **1 Introduction**

In an increasingly digitally connected world, where the deployment of technology continues to expand in all areas of society, changing the way we communicate, work and make business, digitization has become the megatrend for organizations (Klups 2015), acting as an agent of disruptive innovation and leading to the rethinking of business models, processes and strategies to obtain maximum economic benefits and maintain a company’s competitive edge.

Whether as future employees or independent entrepreneurs, students need to be prepared to face such an IT-driven workplace where, according to Anderson et al. (2009), “IT and the business are so intertwined and interconnected that IT is the business, and the business is IT”. Burton (2012), however, regards technology as still “...usually viewed from too much of an IT perspective and not enough from a business integration and user development perspective” originating “...a huge missing link between the human talent development, and the technologies being marketed...”.

It is therefore vital for future business administrators to acquire the essential IT competencies to understand the underlying concepts and the impact of IT strategic implementations leading to

business change and improvement. Only then will they be able to work together with the IT specialists and participate in the modelling and implementation of sustainable business IT solutions.

### 1.1 Background Information and General IT Curriculum

At Offenburg University, Business Administration students follow two information technology modules as illustrated by Figure 1. The first study phase includes the module *Information Technology 1* which consists of the *IT Fundamentals* course in the first semester followed by the *IT-Tools* course in the second semester. In the second study phase students follow the module *Information Technology 2* consisting of the courses *Business Information Systems*, which is largely based on SAP-ERP case studies, and *Business Intelligence*, which integrates the use of SAP BI tools, also through case studies.

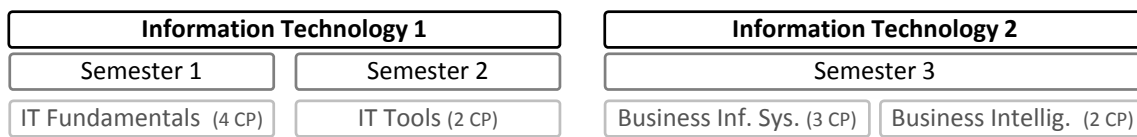


Figure 1: Information technology modules for Business Administration students

### 1.2 Outline of the *IT Fundamentals* Course

The curricular framework of the *IT Fundamentals* course is essentially based on the key IT topics and competencies required by the modern business world, as identified in standard business informatics textbooks such as Abts and Mülder (2013). Before the implementation of the approach described in this article, the course followed the sequence as illustrated in Figure 2. The arrows indicate some of the cross-referencing topics, suggesting that an integrated approach, rather than the traditional chapter sequence, would actually be more suitable for this course.

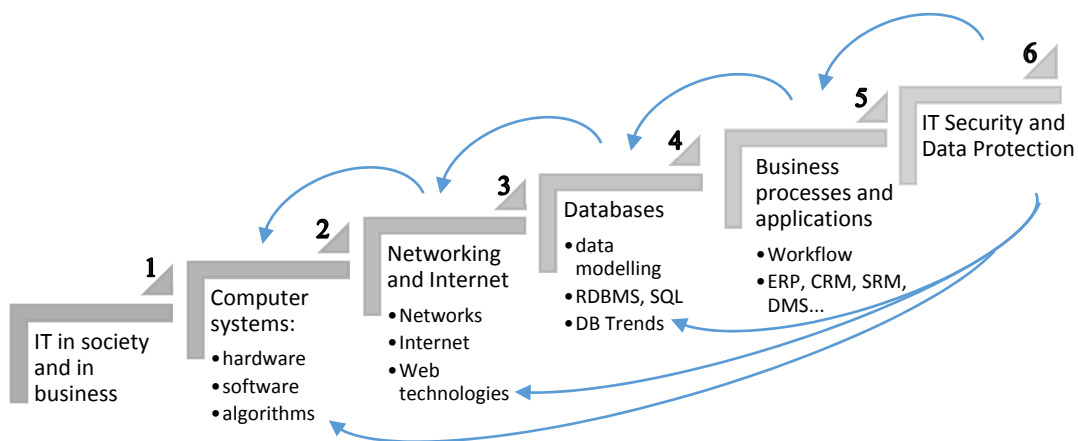


Figure 2: Curriculum framework for the traditional *IT Fundamentals* course

The main problem with this approach was that students learnt the topics under a traditional computer science lecture model, where content proved at times too technical or not relevant in a business context, and with little hands-on experience. The students were therefore more focused on studying to pass the final examination rather than engaging themselves in the learning activities or developing an intrinsic motivation to extend their knowledge. This perception was confirmed by

the students' feedback at the end of the semester and as a result, we decided to explore new ways to teach IT to Business Administration students.

Throughout our research we found that most methods focused on how to teach informatics or on how to teach ERP, but so far no model integrated both. This inspired us to design our innovative concept: the teaching of "IT via ERP".

## 2 ERP and the new Curricular Framework

ERP systems enable the effective integration of business processes, optimizing the operations and the management of data across all enterprise functional areas and are therefore considered the IT backbone for enterprises of virtually any size and in nearly all business fields. However, successful ERP implementation projects tend to be complex and require not only a robust and reliable IT infrastructure but also a thorough definition of business requirements (Anderson et al. 2009).

On analysing the stages of an ERP implementation project, we recognized that all of the IT competence areas, as defined in our curriculum, could be mapped onto an ERP project life cycle, providing the ideal context for an integrated "IT via ERP" model. This new curricular framework was designed based on an adaptation of Leiting's (2012) ERP project structure as illustrated in Figure 3.

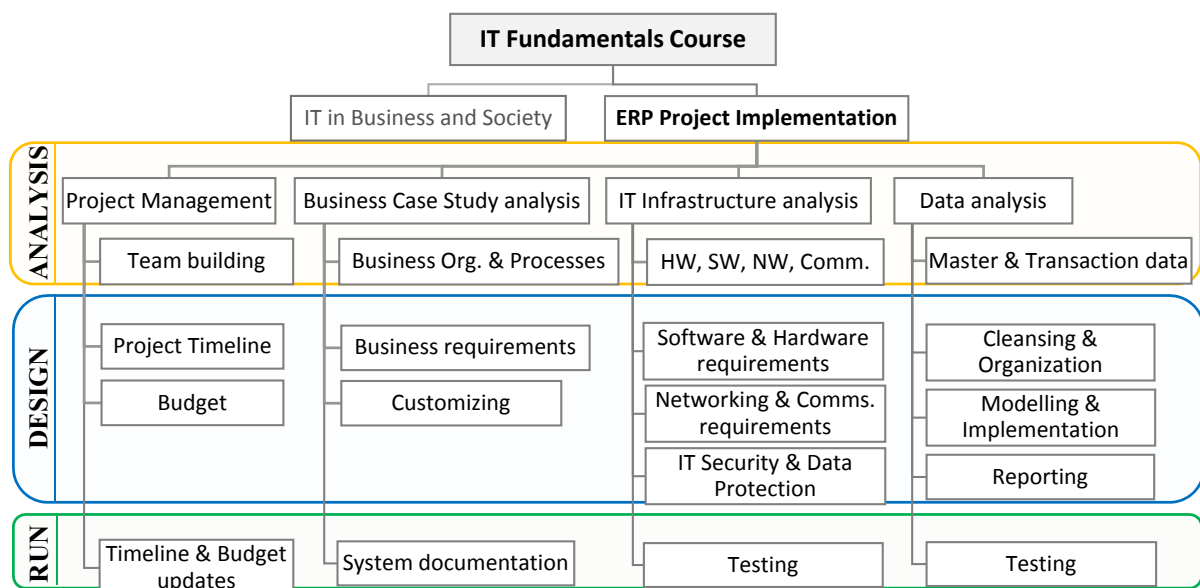


Figure 3: the "IT via ERP" curricular framework

The structure is supported by three tiers – Analysis, Design and Run – which roll across four project areas – Project Management, Business Case Study, IT Infrastructure and Data analysis. The relevant IT topics were matched to each corresponding phase of the ERP implementation project life-cycle. In this way, the learning of IT and the essentials of ERP and business process integration occur simultaneously and students gain a deeper understanding of the importance of an IT-strategy to support the enterprise business objectives.

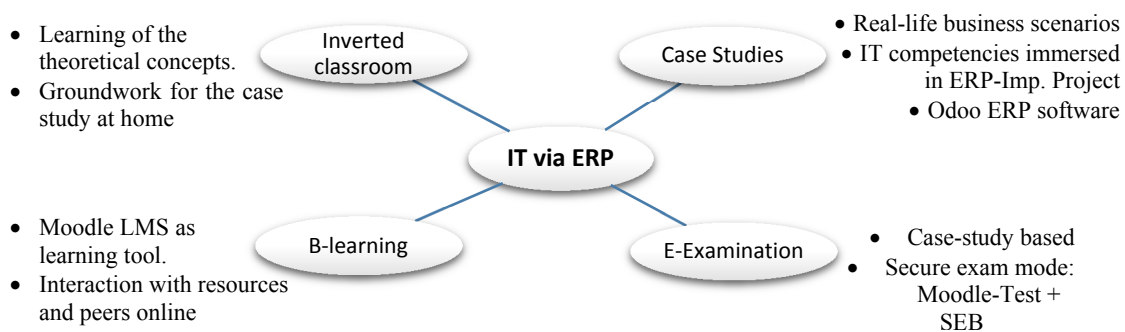
## 2.1 The ERP System

To support the learning activities planned for the proposed framework we chose the open source software Odoo ([www.odoo.com](http://www.odoo.com)), formerly OpenERP, because it provides a simpler and more attractive user interface, where all the functional areas are integrated as one single tool without requiring any additional software. This operational simplicity was taken into account as a factor to impact positively upon the students' learning and motivation curves. As a cloud-based solution, Odoo is straightforward to implement and it is free of charge for educational institutions. Furthermore, Odoo can also be deployed on premise on a standard database and webserver. Its administration and user management is relatively simple, allowing students to work in groups and create different organizations with some customizing options. We also agreed that students would benefit from the exposure to different ERP landscapes throughout their studies, in our case Odoo and SAP, increasing their awareness of the different architectures and process flows within ERP systems (Leyh et al. 2012).

## 2.2 The didactical concept

Through the "IT via ERP" framework we addressed the issue of integrating and contextualising the IT concepts, making them appropriate and relevant for Business Administration students. However, the implementation of such a model also requires a change in pedagogical strategy in order for it to be effective.

The new concept to support this framework consists of four main features: the B-Learning model using Moodle, the inverted classroom model, the case studies approach and an E-Examination as illustrated in Figure 4.



**Figure 4: The didactical concept for "IT via ERP"**

We established a B-Learning model to complement the face-to-face sessions with online activities on Moodle, with the aim of using the LMS as a learning tool and not as a mere repository of resources. We also changed the course structure to suit the new "IT via ERP" framework and published the necessary resources for each topic. These were only made available, though, after completing the previous unit in order to allow students to carry out the tasks in preparation for the next case study. Through the application of this inverted classroom methodology (Langer et al. 2014) we optimized the time in class, promoting an active and collaborative learning environment. The students could also monitor their activity and learning progress in Moodle, helping them to develop a sense of ownership and responsibility, and thus to encourage independent learning.

With the case study approach we simulated a real-life learning scenario in the workplace for our fictitious company – the car accessories dealer "Crash AG". In groups of 4-5, with each group being

an instance of “*Crash AG*”, the students worked collaboratively to build and acquire new knowledge through problem-solving and decision-making activities, helping them also to develop social competencies such as teamwork and interpersonal skills (Barkley et al. 2014).

Table 1 shows the set of case studies designed for a total of 48 semester periods per week (SPW) out of 52, based on each of the ERP-project phases of our framework.

	Project Phase		Case Study	SPW	Odoo
Analysis	IT in Business & Society	1	The Company “Crash AG”: understanding the current business processes and the role of IT	4	---
	ERP Implementation: Project Management	2	ERP-Project Life Cycle: team building, resource planning, timeline and budget	2	- Demo System - PM Module
Design	ERP Implementation: Business Cases	3	Business Case analysis: objectives, workflows, business requirements, proposals, customizing	4	Module analysis and their integration
	ERP Implementation: IT Infrastructure	4	Software: analysis and selection	6	- Module installation - Project update
		5	Hardware: analysis and selection	4	Project update
		6	Networking: analysis and proposals	6	Project update
		7	Web technologies: website design and implementation	6	- Website Module - Project update
		8	IT Security and data protection: analysis and procedures	4	- System user rights - Project update
	ERP Implementation: Data	9	Database implementation: data modelling, customizing, reporting	10	Purchase, Sales, Customers & Logistics Modules
Ru	ERP Implementation: Testing	10	“Crash AG”-ERP goes live!	2	Daily operations simulation game

**Table 1: new Case Study structure under the “IT via ERP” model**

Because many students suffer from what is known as the “bulimic learning” phenomenon and only want to memorize what is to be examined (Haerder 2012), we also changed the assessment format to an E-Examination based on a case study and online Moodle questions, making it consistent with our pedagogical model. Students would then need to demonstrate their IT competences by applying them in a different business context, instead of memorizing and repeating concept definitions.

### 2.2.1 An Example: The Data Analysis Case Study

The unit “Data Analysis” starts with a “flipped activity” at home. Students read the theory notes on Databases published in Moodle. By means of an animation they relate these concepts to a real-life example and then answer the corresponding online quiz in Moodle. In class, the session starts with a short discussion on the “flipped activity” to set the students in context and go over the theoretical concepts needed to solve the case study. Working on the case study in groups, the students:

- Analyse the described processes in the Odoo “Crash AG” system, identifying, extracting and making a list of the actual data and the corresponding data types.

- Organise and normalise the data, identify the primary and foreign key fields, the master and transaction data, and establish the relationships between the different tables.
- Analyse the non-computerised process, as presented in the case study, and design a relational data model for the ERP system.
- Implement the model in MS-Access and create the necessary queries to produce the required reports.

They can use the notes, useful links and other resources provided in Moodle, working under the guidance of the lecturer who takes on the role of a facilitator (Langer et al. 2014). The solutions are presented and discussed upon completion of each section of the case study and at the end the students upload their final work in Moodle for the lecturer's supplementary feedback.

### 3 First results

The evaluation of the application of this didactical model in two semesters generated positive and encouraging results. With an average of fifty students per semester, the level of sustained attendance increased to 90% - 100% in every lesson in comparison with the erratic attendance rates in the past.

The level of motivation and engagement were also measured through the observation of active participation in class, the number of interactions ("clicks") on the activities and resources in Moodle, and the completion of the case studies. As a result, the course was rated as the most active Moodle course in both semesters. Table 2 shows a comparison of results between the previous semester under a traditional approach (\*) and the two semesters with the new methodology, where we also observe a significant improvement in the final examination passing rates.

Semester	Nr. of Students	Total Nr. of Clicks	Activity completion rate	Exam Pass rate
SS 14 (*)	85	7583	10%	88,0%
WS 14/15	51	17480	87%	96,0%
SS 15	45	18361	92%	95,5%

**Table 2: Results after two semesters with the new methodology**

The students had the opportunity to reflect upon their IT competency levels "before and after" by taking part in the online self-assessment surveys at the start and end of the course, and also to comment on the model. The contextualised, integrated and practical nature of the course through the application of the B-Learning approach combined with the case studies was highly valued. They mentioned that they had learned more topics in a relevant context and in a useful way because nowadays it is expected from people to continue to develop their skills online using similar systems. The availability of self-paced tests at the end of each unit was also regarded as very positive.

They mentioned, however, that it took them some time to get used to the "flipped activities" and understand their importance for the case studies. Others commented having a greater workload for this course than for others, feeling at times overwhelmed with the new methodology and the variety of activities. For a few, working in groups proved a challenge at the beginning but this situation improved throughout the semester and under the lecturer's guidance.



## 4 Conclusions

Our curriculum model shows that it is possible to align a standard IT curriculum for Business students along an ERP implementation project. This approach not only covers the relevant IT topics, but also stresses the important role of ERP systems and business process centric thinking for modern enterprises. Through it we were able to increase the levels of engagement and willing attendance to class, leading to improvement in learner achievement. As this way of learning was new to everyone, we also addressed the issue of heterogeneity of individual IT skills and promoted the development of the students' social competencies during the group activities.

Although this model requires a significant initial effort in the planning, designing and development of the resources and the case studies, it was worthwhile to see the change in attitude and levels of motivation of the students as a result of this new approach.

We will continue to refine the model and enhance the case studies based on the semester outcomes and the student feedback. We will also include badges as a gamification element, making each case study a mission or challenge to be accomplished, and in this way motivate our students even further.

## 5 Literature

- Abts D, Müller W (2013) Grundkurs Wirtschaftsinformatik: Eine kompakte und praxisorientierte Einführung. 8th Edition. Springer Fachmedien, Wiesbaden.
- Anderson G, Nilson C, Rhodes T (2009) SAP Implementation Unleashed: A Business and Technical Roadmap to Deploying SAP. 1st Edition. Pearson Education Inc., USA.
- Burton T (2012) Technology: Enabler or inhibitor of improvement? <http://www.processexcellencenetwork.com/business-process-management>. Accessed on 27.07.2015
- Barkley E F, Major C H, Cross K P (2014) Collaborative Learning Techniques: A Handbook for College Faculty. 2<sup>nd</sup> Edition. Jossey-Bass, San Francisco.
- Haerder M (2012) Bildungspolitik Auf den Spuren von Bologna: Das Bulimie Lernen füllt den ganzen Tag aus. <http://www.wiwo.de/politik/deutschland/bildungspolitik-das-bulimie-lernen-fuellt-den-ganzen-tag-aus/6145066-3.html>. Accessed on 10.08.2015
- Langer V, Linke Knut, Schimanke F (2014) Improvement of Self-directed Learning by Using the Inverted Class Model (ICM) for a Basic Module in Business Computer Science. In: Großkurth E, Handke J. (Eds.). The 3<sup>rd</sup> German ICM-Conference - Proceedings: 73-82. De Gruyter.
- Kulps J (2015) Megatrend Digitalisierung: Auf dem Weg in eine neue Arbeitswelt. <https://digitalcologne.de/2015/05/11/megatrend-digitalisierung-auf-dem-weg-in-eine-neue-arbeitswelt/>. Accessed on 05.08.2015
- Leiting A (2012) Unternehmensziel ERP-Einführung: IT muss Nutzen stiften. Springer Fachmedien, Wiesbaden
- Leyh C, Strahringer S, Winkelmann A (2012) Towards Diversity in ERP Education – The Example of an ERP Curriculum. Re-conceptualizing Enterprise Information Systems: Lecture Notes in Business Information Processing Volume 105, 2012: 182-200
- Révészová L (2013) Business Informatics in Managers' Education. In: Management: Science and Education. Volume 2, No. 1: 63-66
- Sora S (2008) The Teaching of Informatics for Business Students. College Teaching Methods & Styles Journal. January 2008, Volume 4, Number 1: 107-110



# Sichere Informationsinfrastrukturen für kleine und mittlere Energieversorger

**Julian Dax<sup>1</sup>, Daniel Hamburg<sup>2</sup>, Michael Kreuzsch<sup>3</sup>, Benedikt Ley<sup>1</sup>, Sebastian Pape<sup>4</sup>, Volkmar Pipek<sup>1</sup>, Kai Rannenber<sup>4</sup>, Christopher Schmitz<sup>4</sup> und Frank Terhaag<sup>5</sup>**

<sup>1</sup> Universität Siegen, Institut für Wirtschaftsinformatik, {vorname.nachname}@uni-siegen.de

<sup>2</sup> TÜV Rheinland i-sec GmbH, Köln, daniel.hamburg@i-sec.tuv.com

<sup>3</sup> Arbeitsgemeinschaft für sparsame Energie- und Wasserverwendung, Köln, kreusch@asew.de

<sup>4</sup> Goethe-Universität Frankfurt am Main, M-Chair, {vorname.nachname}@m-chair.de

<sup>5</sup> regio iT gesellschaft für informationstechnologie mbh, Aachen, frank.terhaag@regioit.de

## Abstract

Eine sicher funktionierende Energieinfrastruktur ist für fast alle Lebensbereiche unserer heutigen Gesellschaft grundlegend. Damit die Energieversorgung im Rahmen der Energiewende auch nachhaltig sichergestellt werden kann, wird auch im Energiesektor immer mehr Informations- und Kommunikationstechnik (IKT) eingesetzt. Gleichzeitig erhöht sich dadurch jedoch auch das Risiko für IT-basierte Angriffe auf die Kritische Infrastruktur. Die effektive und effiziente Absicherung der eigenen Infrastruktur vor solchen Angriffen stellt insbesondere kleine und mittelgroße Energienetzbetreiber mit begrenzten Ressourcen vor eine besondere Herausforderung. Im Forschungsprojekt SIDATE (<http://www.sidate.org/>) setzen wir an dieser Stelle an und erarbeiten geeignete Konzepte und Werkzeuge, die kleine und mittelgroße Energieversorger bei der Verbesserung ihrer IT-Sicherheit angemessen unterstützen sollen. Der Forschungsschwerpunkt liegt dabei auf den drei Kernthemen Bewertung von IT-Sicherheit, Wissensaustausch und Zertifizierung. Wir verfolgen im Projekt eine praxisnahe Forschungsmethodik die eine intensive Einbeziehung der Energieversorger in den gesamten Forschungsprozess vorsieht.

## 1 Motivation

Kritische Infrastrukturen spielen eine wichtige Rolle für das Funktionieren heutiger Informationsgesellschaften. Der Schutz dieser Infrastrukturen liegt dementsprechend im Interesse der Allgemeinheit. Faktoren wie die zunehmende Komplexität der Infrastrukturen, die wachsende Abhängigkeit zwischen unterschiedlichen Infrastrukturen und der vermehrte Einsatz von (mobiler) Informations- und Kommunikationstechnik (IKT) stellen hierbei eine besondere Herausforderung bei der Erreichung dieses Zieles dar. Weiterhin wird gerade im Energiesektor eine starke Änderung der Geschäftsmodelle erwartet (PwC 2014). Diese erfordert in Verbindung mit dem Einsatz von IKT nicht nur die weiter zunehmende Dezentralisierung kritischer Infrastrukturen sondern auch die

Automatisierung von Abläufen zur Steigerung der Effektivität und Effizienz. Dies führt gleichzeitig aber auch zu neuen Risiken und Gefahren.

Die Betreiber im Energiesektor stehen nun vor der Problemstellung, sowohl den Schutz als auch die Wirtschaftlichkeit ihrer Infrastrukturen sicherzustellen. Dieses Spannungsfeld stellt besonders kleine und mittelgroße Energieversorger (z.B. Stadtwerke) vor neue Schwierigkeiten, die ohne fremde Unterstützung kaum zu lösen sind. Für diese Unternehmen ist es in der Regel wirtschaftlich nicht darstellbar, spezialisiertes Personal zum Schutz ihrer IKT-Infrastruktur zu beschäftigen. Auch die Anwendung und Adaption bereits existierender Kriterienkataloge zur IT-Sicherheit bspw. auf Grundlage der ISO/IEC 27001 und ISO/IEC 27002, sind insbesondere ohne spezifische Kenntnisse im Bereich IT-Sicherheit kaum durchführbar. Für die konkrete Umsetzung sicherheitstechnischer Maßnahmen wäre der Einsatz von Dienstleistern eine gangbare Alternative zum Aufbau eigener Ressourcen. Es ist jedoch im Interesse der Energieversorger, den grundlegenden Schutzbedarf und das Sicherheitsniveau ihrer IKT-Infrastruktur eigenständig bzw. im organisationsübergreifenden Wissens- und Erfahrungsaustausch mit anderen Energieversorgern zu ermitteln und gegebenenfalls zu erhöhen.

Das Projekt setzt sich zum Ziel, kleinen bis mittleren Betreibern von kritischen Infrastrukturen Plattformen, Werkzeuge und Modelle zur Verfügung zu stellen, die dabei helfen, selbstständig Schwachstellen zu identifizieren, geeignete Lösungsansätze zu entwickeln und effizienten Gebrauch von organisationsübergreifend vorhandenem Wissen zu machen. Dabei geht es nicht nur um Fragen technischer, sondern auch organisatorischer und prozessualer Sicherungskonzepte. Betrachtet werden dabei sowohl Sicherheitsaspekte (Ausfallrisiko, Angriffsrisiko, Schadensabwendung und -begrenzung) als auch Kooperations- und Usability-Aspekte (z.B. Aufwand, Beeinträchtigungspotenzial und Akzeptanz von Sicherheitsmaßnahmen). Personal mit allgemeinen IKT-Kenntnissen soll in die Lage versetzt werden, den Überblick über Gefährdungen zu behalten und gegebenenfalls durch das Hinzuziehen externer Experten nachhaltige Lösungen zu erarbeiten, die auch im Alltag der Organisation gelebt werden können.

## **2 Forschungs- und Anwendungsfeld**

Der deutsche Energiemarkt zeichnet sich durch eine große Vielfalt aus. Mit der Liberalisierung des Strommarkts Ende der 90er Jahre hat sich die Zahl der Wettbewerber in diesem Bereich stark erhöht. Im April 2015 waren in Deutschland fast 1.200 Stromlieferanten und über 900 Stromnetzbetreiber tätig (BDEW 2014). Insbesondere das Verteilnetz wird dabei größtenteils von meist kleinen bis mittelgroßen kommunalen Energieversorgern (Stadtwerken) betrieben.

Durch den vermehrten Einsatz von IKT, insbesondere auch im Rahmen der Energiewende (z.B. durch eine intelligente und dezentralisierte Energienetzführung), sind die Versorgungsnetze in ihrer IKT-Vernetzung immer weniger autark und es eröffnen sich neue Möglichkeiten IKT-basierter Angriffe auf die kritische Versorgungsinfrastruktur (BNetzA 2011). Ein medienwirksamer Selbsttest der Stadtwerke Ettlingen hat deutlich gemacht, dass die notwendige Absicherung dieser Infrastruktur insbesondere bei kleinen Energieversorgern aufgrund mangelnder Ressourcen und somit fehlender Expertise unzureichend sein kann (Grefe 2014).

Mit der Verabschiedung des IT-Sicherheitsgesetzes sowie der Veröffentlichung des IT-Sicherheitskatalogs für Energieversorger der Bundesnetzagentur im Sommer 2015 wurde ein gesetzlicher Rahmen geschaffen, der die Energieversorger zur Umsetzung und Einhaltung vorgegebener Mindeststandards verpflichtet. Auch kleine Energieversorger, die über wenige

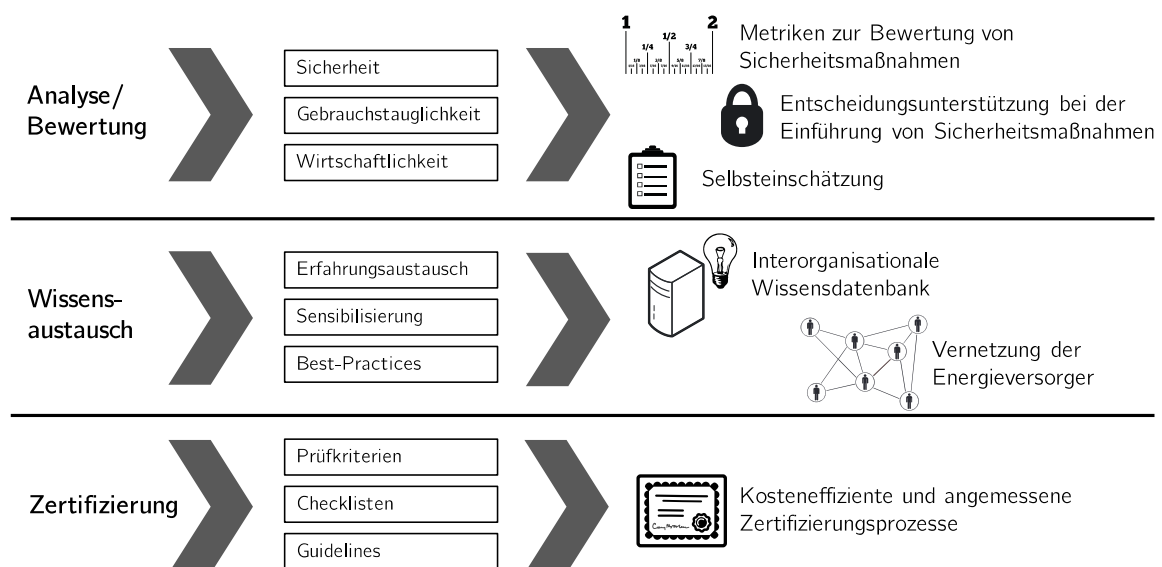
Ressourcen verfügen, müssen die Vorgaben des IT-Sicherheitskataloges uneingeschränkt umsetzen.

Im Forschungsprojekt wollen wir an dieser Stelle ansetzen und geeignete Konzepte entwickeln, die kleine und mittelgroße Energieversorger bei Erfüllung der gesetzlichen Vorgaben und somit bei der Verbesserung ihrer IT-Sicherheit unterstützen. Um dieses Ziel sinnvoll erreichen zu können, arbeiten wir im Projekt mit mehreren kleinen und mittelgroßen Stadtwerken zusammen, die in ihrer Organisation und Infrastruktur unterschiedlich aufgestellt sind (vgl. Tabelle 1).

	Stadtwerk A	Stadtwerk B	Stadtwerk C	Stadtwerk D	Stadtwerk E
Mitarbeiter	1.500	655	680	245	207
stromnetzrelevante Daten:					
Gesamtlänge Freileitungen (km)	541	342	260	200	59
Gesamtlänge Kabel (km)	14.640	6.250	2.192	2.884	799
Netzebenen	Höchst-, Hoch-, Mittel-, Niederspannung	Hoch-, Mittel-, Niederspannung	Hoch-, Mittel-, Niederspannung	Hoch-, Mittel-, Niederspannung	Mittel-, Niederspannung
Versorgte Einwohner (Niederspannung)	428.794	278.373	252.221	208.408	76.628

**Tabelle 1: Übersicht assoziierter Energieversorger**

Die Zusammenarbeit mit und Gegenüberstellung von verschiedenen Energieversorgern im Forschungsvorhaben ermöglicht es uns, Generalisierungen auf Basis einer einzelnen Organisation zu vermeiden. Für jedes beteiligte Unternehmen werden daher basierend auf den existierenden technischen und organisatorischen Begebenheiten individuelle Praktiken und Prozesse zur Umsetzung und Verbesserung von IT-Sicherheit sowie damit einhergehende Probleme untersucht.



**Abbildung 1: Übersicht des geplanten Forschungsvorhabens**

### 3 Forschungsfokus und -fragen

Im Rahmen des Projekts sollen verschiedene Forschungsbeiträge entstehen. Diese konzentrieren sich im Wesentlichen auf die drei übergeordneten Kernthemen Bewertung von IT-Sicherheit, Wissensaustausch und Zertifizierung, welche in Abbildung 1 skizziert werden (vgl. Abb. 1).

Als Ausgangsbasis weiterer Forschungsschritte werden zunächst IT-sicherheitsrelevante Prozesse sowie Sicherheitsrisiken kleiner und mittlerer Energieversorger identifiziert. Darauf aufbauend werden Angriffsszenarien erarbeitet und ein Maßnahmenkatalog zur Absicherung der IKT-Infrastruktur abgeleitet.

Als zentraler Aspekt der Bewertung von IT-Sicherheit ist der Entwurf leichtgewichtiger Metriken und zugehöriger Erhebungsprozesse hervorzuheben, anhand derer das IT-Sicherheitsniveau gemessen und konkrete Sicherheitsmaßnahmen hinsichtlich ihrer Sicherheit, Gebrauchstauglichkeit und Wirtschaftlichkeit beschrieben werden können. Leichtgewichtigkeit meint, dass die Metriken auch ohne besonderen personellen oder finanziellen Ressourcenaufwand direkt durch kleine und mittlere Energieversorger anwendbar sind. Hierdurch sollen die Sicherheit, Gebrauchstauglichkeit und Wirtschaftlichkeit auch für Nicht-Experten einfach und effizient messbar werden. Dies ermöglicht eine regelmäßige Selbsteinschätzung des eigenen IKT-Sicherheitsniveaus sowie möglicher Sicherheitsmaßnahmen. Der Entwurf der Metriken, wie auch der sonstigen im Projekt entwickelten Werkzeuge, erfolgt dabei von Beginn an unter Gesichtspunkten der Nachhaltigkeit. So sollen alle entwickelten Technologien leicht an neue Technologien (z.B. Smart Meter) angepasst werden können und so nicht nur vorhandene, sondern auch sich durch Veränderungen ergebende, Systeme betrachtet werden.

1. Die quantitative Beschreibung von IT-Sicherheit stellt eine der primären Forschungsaufgaben dar. Es sollen aussagekräftige Metriken entworfen werden, um das IT-Sicherheitsniveau und konkrete Sicherheitsmaßnahmen adäquat bewerten zu können. Dies umfasst sowohl die rein technische Ebene zur Bewertung von IT-Infrastrukturen als auch die organisationale Prozessebene, um eine gesamtheitliche Perspektive auf die IT-Sicherheit einzunehmen.

Generell lässt sich IT-Sicherheit nur schwer messen (Pfleeger 2010). Die besondere Herausforderung im Rahmen dieses Projekts besteht zusätzlich darin, Metriken zu entwickeln, die einerseits über genügend Aussagekraft besitzen, andererseits aber auch leichtgewichtig sind. Es gilt also, einen Trade-off zwischen Einfachheit der Anwendung und Genauigkeit der Ergebnisse zu finden.

2. Bedienbarkeit, Verständlichkeit und Akzeptanz von Sicherheitsmaßnahmen haben einen wesentlichen Einfluss auf das sicherheitsrelevante Nutzerverhalten. So werden Sicherheitsmechanismen, die nicht verstanden oder als störend empfunden werden, häufig durch unbewusste Fehlbedienung oder bewusstes Aushebeln wirkungslos (Whitten und Tygar 1999, Herzog 2007). Dennoch sind existierende Entwicklungsprozesse und Vorgehensmodelle des Security-Engineerings von denen des Usability-Engineerings noch immer weitgehend entkoppelt (Fischer-Hübner et al. 2011). IT-Sicherheitsverantwortliche in Unternehmen entscheiden zudem über Sicherheitsrichtlinien und -mechanismen, ohne sich der damit einhergehenden Auswirkungen auf die Arbeitspraxis der Mitarbeiter bewusst zu sein (Parkin et al. 2013).

Eine wesentliche Forschungsaufgabe im Projekt wird daher sein, solche Auswirkungen sichtbar und messbar zu machen, sodass diese bei der Auswahl geeigneter Sicherheitsmaßnahmen berücksichtigt werden können.

3. Insbesondere für kleine und mittlere Energieversorger nimmt die Wirtschaftlichkeit von Sicherheitsmaßnahmen aufgrund ihrer geringeren finanziellen Ressourcenausstattung einen hohen Stellenwert ein. Das Anspruchsvolle an der ökonomischen Bewertung von Sicherheitsmaßnahmen ist, dass durch IT-Sicherheit im Allgemeinen keine Umsätze generiert werden, sondern nur Schaden reduziert wird, der sich nicht exakt quantifizieren lässt.

Im Projekt soll die Wirtschaftlichkeit anhand der für die Durchführung von Sicherheitsmaßnahmen anfallenden Kosten sowie deren Wirksamkeit ermittelt werden. Hierfür werden individuelle Kosten- und Wirksamkeitsfaktoren für die identifizierten Sicherheitsmaßnahmen abgeschätzt und daraus Effizienzfaktoren abgeleitet.

Darüber hinaus wird ein Konzept zur Visualisierung von Infrastrukturelementen erarbeitet, welches ihre Zusammenhänge mit Organisationseinheiten und Geschäftsprozessen darstellt und ergänzende Informationen zu IT-Sicherheitsniveau, Usability und Kosten enthält. Letztendlich sollen Entscheider hierdurch virtuell Sicherheitsmaßnahmen einführen können und die verschiedenen Auswirkungen grafisch angezeigt bekommen.

Neben der Bewertung von Sicherheitsmaßnahmen kommt dem Wissensaustausch eine essentielle Rolle zu. So soll eine organisationsübergreifende Wissensdatenbank und Austauschplattform zum Erfahrungsaustausch mit anderen Energieversorgern entstehen. Dies ist insbesondere deshalb sinnvoll, da Infrastrukturen und Risikostrukturen von Betreibern kritischer Infrastrukturen oft vergleichbar sind und Lösungsansätze somit sinnvoll wiederverwendet werden können.

In Hinblick auf die Etablierung einer solchen Plattform ist zunächst die Nutzungsbereitschaft der Energieversorger genauer zu betrachten, die sich im Sinne einer Kosten-Nutzen-Abwägung aus Nutzungsmotivation einerseits und Nutzungsaufwand und -risiken andererseits ableitet.

So stellt sich etwa die Frage, welcher Nutzungsaufwand in Bezug auf Datenerhebung und -erfassung vertretbar ist und welche Informationen über die IT-Infrastruktur, dessen Sicherheitsniveau und eingeführte Maßnahmen und Auswirkungen in der Wissensdatenbank abgelegt werden sollen. Dies umfasst auch die Entwicklung adäquater Formate und Datenstrukturen für die Speicherung relevanter Informationen sowie der Identifizierung möglicher Assoziationen zwischen verschiedenen Informationstypen.

Da die Energieversorger über die Plattform höchst sensible Informationen austauschen, sind adäquate Datenschutz- und Zugriffskontrollkonzepte zu entwickeln. Diese sollen sicherstellen, dass die Energieversorger die Kontrolle darüber haben, welche Informationen in welchem Detailgrad abgespeichert werden und wem diese Informationen zur Verfügung gestellt werden.

Darüber hinaus werden Zertifizierungskonzepte für Energieversorger entwickelt. Dafür werden speziell an die Bedürfnisse kleiner und mittlerer Energieversorger angepasste Prüfkriterien, Checklisten und Guidelines erarbeitet, um eine kosteneffiziente Zertifizierung gemäß gültiger Standards und gesetzlicher Vorgaben zu ermöglichen. Mit einer solchen Zertifizierung sollen Energieversorger dem Gesetzgeber, aber auch ihren Endkunden und Geschäftspartnern gegenüber, die Einhaltung eines hohen Sicherheitsniveaus nachweisen können. Gleichzeitig wäre dies ein Beitrag zur Erhöhung der Markttransparenz für Endkunden.

Es soll ein Anforderungskatalog auf Basis der entwickelten Metriken konzipiert werden, der es erlaubt, eine möglichst vollständige Auflistung wichtiger Anforderungen zu identifizieren, die benötigt werden, um eine Zertifizierungsaussage treffen zu können. Dieser orientiert sich an zuvor ausgearbeiteten relevanten Prüfkriterien, die im Rahmen einer Zertifizierung zu prüfen sind. Die Herausforderung bei der Erstellung eines solchen Prüfkatalogs besteht im Ausarbeiten eindeutiger und möglichst objektiv zu bewertender Prüffragen. Bei der Ausarbeitung des Prüfkatalogs sollen auch bestehende relevante Zertifizierungen berücksichtigt werden. Die Berücksichtigung bestehender Zertifizierungen verfolgt das Ziel, die Kosten für die weitere Zertifizierung zu minimieren. Hierzu muss die Abdeckung der Prüfpunkte von vorhandenen Zertifizierungen analysiert werden. Je nach Abdeckungsgrad der Anforderungen können zu prüfende Punkte aus einer bestehenden Zertifizierung abgeleitet werden und somit der Prüfumfang verringert werden.

#### **4 Methodisches Vorgehen**

Das Forschungsvorhaben sieht eine praxisnahe Forschungsmethodik bestehend aus (1) einer Analyse der Arbeitspraxis, (2) der Entwicklung von Konzepten und Softwarewerkzeugen und (3) deren Evaluation in der Praxis vor (Wulf et al. 2015). Ziel der ersten Phase ist es, mittels qualitativer und quantitativer empirischer Forschungsmethoden die derzeitige Arbeits- und Kooperationspraxis bei Energieversorgern hinsichtlich systemsicherheitsrelevanter Prozesse zu untersuchen, um Anforderungen an mögliche Konzepte und Werkzeuge zur Unterstützung dieser Praxis zu erhalten. Dazu werden verschiedenen Studien bei assoziierten Energieversorgern vor Ort im Rahmen von Interviews, Workshops, Beobachtungen und Dokumentenanalysen durchgeführt. Außerdem wird gemeinsam mit einem deutschlandweit agierenden Stadtwerkenetzwerk ein regelmäßig stattfindender Arbeitskreis zum Thema IT-Sicherheit für Energieversorger gebildet, der den bilateralen Austausch und die Kooperation zwischen dem Projektkonsortium und interessierten Stadtwerken ermöglichen soll.

Bei der Entwicklung der Konzepte und Werkzeuge verfolgen wir einen partizipativen und iterativen Ansatz. Die Grundlage hierfür bilden die Ergebnisse der Praxisanalyse sowie relevante rechtliche und regulatorische Rahmenbedingungen. Die (Zwischen-)Ergebnisse werden in regelmäßigen Abständen im Rahmen von Workshops, Fokusgruppen sowie dem erwähnten Arbeitskreis mit Anwendern diskutiert, validiert und weiterentwickelt.

Ein wesentliches Ziel ist es, die Konzepte und Werkzeug-Demonstratoren so zu gestalten und zu entwickeln, dass diese im Rahmen des Forschungsprojekts im Anwendungsfeld unter realen Bedingungen evaluiert und auf ihre Praxistauglichkeit hin untersucht werden können.

#### **5 Zusammenfassung**

In dieser Arbeit wurde aufgezeigt, dass kleine und mittlere Energieversorger sowohl den Schutz als auch insbesondere die Wirtschaftlichkeit ihrer Infrastrukturen sicherstellen müssen. Dabei sind die Energieversorger zur Umsetzung und Einhaltung vorgegebener Mindeststandards durch die Vorgaben des IT-Sicherheitsgesetzes verpflichtet. Das Projekt hat zum Ziel, die Energieversorger dabei zu unterstützen, insbesondere in drei übergeordneten Kernthemen: der Bewertung von IT-Sicherheit, Wissensaustausch zwischen und Zertifizierung der Energieversorger. Die Erfassung der Anforderungen, die Entwicklung der Konzepte und Demonstratoren sollen dabei in enger Kooperation mit den assoziierten Partnern, kleinen und mittelgroßen Energieunternehmen, als



iterativer Prozess durchgeführt werden. Besonderes Augenmerk liegt dabei auch auf der Zukunftsfähigkeit der entwickelten Konzepte, das heißt diese sollen möglichst leicht an neue Technologien angepasst werden können.

*Danksagung:*

Die Forschungsarbeiten werden durch das Bundesministerium für Bildung und Forschung (BMBF) im Rahmen des Förderschwerpunktes „IT-Sicherheit für Kritische Infrastrukturen“ gefördert.

## 6 Literatur

- BDEW (2015) Marktteilnehmer Energie aktuell - Vielfalt im Energiemarkt. [https://www.bdew.de/internet.nsf/id/5512898B85FDC9C1C12579C2004225A8/\\$file/Marktteilnehmer%20Energie%20aktuell\\_online\\_o\\_halbjaehrlich\\_Ki\\_10042015.pdf](https://www.bdew.de/internet.nsf/id/5512898B85FDC9C1C12579C2004225A8/$file/Marktteilnehmer%20Energie%20aktuell_online_o_halbjaehrlich_Ki_10042015.pdf). Abgerufen am 18.09.2015
- BNetzA (2011) „Smart Grid“ und „Smart Market“ - Eckpunktepapier der Bundesnetzagentur zu den Aspekten des sich verändernden Energieversorgungssystems.
- Fischer-Hübner S, Grimm R, Lo I, Möller S, Müller G, Volkamer M (2011) Gebrauchstaugliche Informationssicherheit. Die Zeitschrift für Informationssicherheit Jg. 2011 (4):14-19
- Grefe C (2014) „Blackout“. In: Die Zeit, Nr. 16/2014
- Herzog A (2007) Usable Security Policies for Runtime Environments. *Linköping Studies in Science and Technology*. PhD Thesis, Department of Computer and Information Science, Linköping University
- ISO/IEC 27001 (2015) IT-Sicherheitsverfahren – Informationssicherheits-Managementsysteme – Anforderungen
- ISO/IEC 27002 (2008) IT-Sicherheitsverfahren – Leitfaden für das Informationssicherheits-Management
- Parkin S, van Moorsel A, Ingelsant P (2010) A stealth approach to usable security: helping IT security managers to identify workable security solutions. In: Workshop on New Security Paradigmns (NSPW '10). 33-50.
- PwC (2014) 14th PwC Global Power & Utilities Survey: "A different energy future -- Where energy transformation is leading us". PricewaterhouseCoopers International Limited.
- Pfleeger SL, Cunningham RK, "Why Measuring Security is hard". IEEE Security & Privacy 8(4):46-54
- Whitten A, Tygar, JD (1999) Why Johnny can't encrypt: a usability evaluation of PGP 5.0. In: Proceedings of the 8th conference on USENIX Security Symposium.
- Wulf V, Müller C, Pipek V, Randall D, Rohde M, Stevens G (2015) Practice-based Computing: Empirically-grounded Conceptualizations derived from Design Case Studies. In: Wulf V, Schmidt K, Randall D (Hrsg.) Designing Socially Embedded Technologies in the Real-World, Springer, London, 111-150



# Erhebungsinstrument zur Messung der Nutzungsmotivation von Energieverbrauchsanzeigen im Automobil

Matthias Gottlieb<sup>1</sup>, Mateusz Burek<sup>1</sup>, Markus Böhm<sup>1</sup> und Helmut Krcmar<sup>1</sup>

<sup>1</sup> Technische Universität München, Lehrstuhl für Wirtschaftsinformatik,  
{gottlieb|burek|markus.boehm|krcmar}@in.tum.de

## Abstract

Eine Verringerung der durch den Straßenverkehr verursachten Emissionen kann durch Beeinflussung des Verhaltens der Fahrer erreicht werden. Dabei kann Informationstechnologie genutzt werden, um den Fahrer zu einem energiebewussten Fahren zu motivieren. Die Motivationsforschung zeigt eine Vielzahl an Theorien zur Erklärung von Verhalten auf. Jedoch ist nicht bekannt, ob Energieverbrauchsanzeigen zur Nutzungsmotivation im Fahrzeug und somit zum effizienteren Fahren beitragen. Als Zwischenschritt wurde (i) eine Literaturrecherche zur Identifikation von motivationalen Einflussfaktoren durchgeführt, und (ii) ein erstes Erhebungsinstrument zur Messung der Nutzungsmotivation von Energieverbrauchsanzeigen im Automobil entwickelt. Ein erstes Zwischenergebnis zeigt, dass kein ausreichendes Energiebewusstsein (EBS) besteht und Energiesparen von eigenen, hauptsächlich intrinsisch geprägten Werten abhängt. Eine vollständige Validierung des Erhebungsinstruments ist notwendig und es ist bisher unbekannt, welcher Interaktionskanal genutzt werden kann, um das EBS zu erhöhen.

## 1 Motivation

Der weltweite Energiebedarf steigt kontinuierlich an und die atmosphärische Treibhausgaskonzentration erreichte im Jahr 2013 einen neuen Höchstwert (WMO 2014). Diese Veränderung ist zurückzuführen auf menschliches Handeln wie die Verbrennung von fossilen Rohstoffen durch Kraftfahrzeuge (WMO 2014; National-Research-Council 2010; IPCC 2007). Verhalten wird durch Motivation gesteuert. Motivation ist die „Energie und Richtung“ einer Handlung (Elliot und Covington 2001, 73). Dies bedeutet Bedürfnisse und Verlangen sind der Auslöser für das Verhalten inklusive der Fahrzeugsteuerung. Es kann einen persönlichen, intrinsischen Grund für eine Handlung (Basis ist Interesse) oder äußeren, extrinsischen Anreiz geben (Basis ist Belohnung oder Bestrafung). Informationstechnologie (IT) als externer Faktor ermöglicht es, menschliche Motivation zum energiebewussteren Fahren zu beeinflussen. Trotz CO<sub>2</sub>-Emissionsgrenzwerten und effizienten Autos bleibt der Mensch die signifikante Kraft die zwischen hohen und niedrigen Emissionsausstoß mit ihrem Verhalten entscheidet. Die Motivationsforschung zeigt in den letzten Jahrzehnten eine Zunahme an Theorien zur Individualisierung und Bestimmung von Einflussfaktoren auf Hand-

lungsweisen auf (James 1890; Skinner 1938; Hull 1943; Maslow 1943; McClelland 1961; Vroom 1964; Locke 1965; Deci 1971; Deci und Ryan 1985; Ajzen 1991; Barbuto und Scholl 1998; Bandura 1999; Venkatesh und Davis 2000; Levesque et al. 2010; Looock et al. 2013). Jedoch ist die Korrelation zwischen Motivation zum Energiesparen und EBS beim Autofahren wenig erforscht. Diese Motivation kann mittels IT beeinflusst werden. Es ist wenig bekannt, welche Interaktionskanäle zu einem stärkeren EBS führen und wie es gemessen werden kann. Dies bedeutet, die Nutzung von soziotechnischen Systemen aus dem Bereich der Mensch-Maschine-Interaktion zur Beeinflussung des menschlichen Verhaltens hin zu einem energiebewussteren Verhalten. Diese Systeme könnten in einem weiteren Schritt durch Experimente näher betrachtet werden. Der Praxis ermöglicht es kundenorientierte Energieverbrauchsanzeigen für ein geeignetes Feedback zu gestalten.

## 2 Methodik und Ergebnisse der Literaturrecherche

Es wird eine Literaturrecherche nach Brocke und Simons (2009), für einen Überblick zur Messung von Motivation genutzten Einflussfaktoren und Methoden, durchgeführt. Es werden die Datenbanken: *Google Scholar*, *EBSCOhost* und *ScienceDirect* verwendet und die ersten 200 Treffer werden nach Titel und Abstrakt auf Relevanz geprüft. Die verwendeten Schlüsselwörter sind: *motiv\**, *intrinsic\**, *extrinsic\**, *car\**, *influence\**, *energy\**, *behavior\** und *predict\**. Zunächst werden die motivationalen Einflussfaktoren aus der Literatur abgeleitet. Danach kann die Messung der Motivation genauer betrachtet werden. Die aufgeführte Literatur beschränkt sich aus Platzgründen auf 31 rezitierte Quellen (28 Theorien, 3 Fragebögen), die für das Erhebungsinstrument als relevante Literatur angesehen werden. Diese werden in einem Fragebogen zusammengefasst und mit alltäglichen Fragen zum EBS validiert.

### 2.1 Motivationale Einflussfaktoren

Motivationsforschung kann im Kern in zwei Bereiche aufgeteilt werden: Inhalts- und Prozesstheorien. Während die Inhaltstheorien die Bedürfnisse des Menschen und deren Veränderungen im Laufe der Zeit erforschen, beschäftigen sich die Prozesstheorien mit der Entstehung der Motivation. Beide Modelle basieren auf dem inzwischen verworfenen instinktiven Denkansatz (James 1890; McDougall 1908), der Triebreduktions- und Anreiztheorie (Woodworth 1917; Hull 1943; Watson 1913; Skinner 1938) sowie der Erregungstheorie (Harlow 1953; Hebb 1955; Zuckerman 1979). Zudem gibt es intrinsische und extrinsische motivationale Einflussfaktoren. Die Literaturrecherche zu motivationalen Einflussfaktoren beinhaltete die Aufarbeitung von 200 Journalen, 45 Büchern, sechs Konferenzen und 26 Webseiten. Die durchschnittliche Relevanz der geprüften Werke gemessen an der subjektiven Eignung, welche mittels einer Skala zwischen 0 (nicht relevant) bis 5 (sehr relevant) ermittelt wurde, beträgt 3,98 Punkte. Die Einteilung in eine Skala erfolgte zusätzlich anhand der Schlüsselwörter und dessen Vorkommen in Titel, Abstract und Text.

**Inhaltstheorien:** Maslow (1954) entdeckte fünf hierarchisch aufgebaute Bedürfnisse, die jeder Mensch in sich trägt. Neben den physiologischen Grundbedürfnissen sind die Sicherheits- und Zugehörigkeitsbedürfnisse sowie das Bedürfnis nach Wertschätzung, Selbstverwirklichung und Selbsttranszendenz zu betrachten. McClelland (1961) fasst diese Faktoren in den Bedürfnissen Macht, Leistung und Zugehörigkeit zusammen. Alderfer (1969) hingegen sieht die Grundbedürfnisse der Menschen in dem Bestreben nach einer sicheren Existenz, einer stabilen Beziehung und der Möglichkeit sowohl beruflich als auch persönlich zu wachsen. Deci und Ryan (1985) publizieren ihre Selbstbestimmungstheorie mit der Behauptung, dass das Gefühl kompetent zu sein und

autonom zu handeln, maßgeblich an der Motivation des menschlichen Handelns beteiligt ist. Außerdem sei das soziale Eingebundenheitsgefühl ein wichtiger Faktor zur Steuerung der Motivationsstärke.

**Prozesstheorien:** Adams (1963) zeigte, dass das Gerechtigkeitsgefühl maßgeblich an der Entwicklung des Verhaltens beteiligt ist. Vroom (1964) und Porter mit Lawler (1968) versuchten Motivation als das multiplizierte Ergebnis aus der Instrumentalität, der Valenz und der Erwartung, dass ein Verhalten wirkungsvoll ist zu messen. Locke (1968) hingegen zeigt, dass ambitionierte und spezifizierte Ziele zur hohen Motivation führen. Spätere Untersuchungen (Seligman und Darley 1977; Becker 1978) stützen Lockes Behauptung und stellen fest, dass Feedback auf hoch gesteckte Ziele die Motivation zusätzlich fördern. Basierend auf Ajzen und Fischbeins Theorie des überlegten Handelns (1975) verfassten Davis et al. (1989) das Technologie Akzeptanz Model (TAM) mit der Intention herauszufinden, welche Faktoren den Menschen zur Nutzung von Technologie motivieren. Dieses wurde erweitert durch TAM2, UTAUT (Vereinheitlichung dieser Modelle) und TAM3.

## 2.2 Messung von Motivation

Aus 160 wissenschaftlichen Arbeiten lässt sich ableiten, dass Motivation experimentell oder mithilfe von Fragebögen gemessen werden kann. Experimentelle Messungen zeigen meistens den Zusammenhang zwischen extrinsischer und intrinsischer Motivation auf. Deci führte Experimente durch, um festzustellen ob extrinsische Belohnungen die intrinsische Motivation beeinflussen. Das Ergebnis zeigte, dass Belohnungen kurzzeitig die Leistung von freiwilligen Arbeiten erhöhen, jedoch unmittelbar danach ein Abfall von interner Motivation festzustellen war (Deci 1971, 1972b, 1972a). In der Literatur wurden 17 verschiedene Messinstrumente zur Messung von Motivation gefunden. Diese lassen sich in vier Kategorien einordnen: Arbeitsmotivation, Lernmotivation, Glaubensmotivation (Religion) und die der Akzeptanz von Technologie. Barbuto und Scholl (1998) stellten eine Zusammenfassung von üblichen motivationalen Messmethoden basierend auf 14 verschiedenen Motivationstheorien vor. 13 weitere Messinstrumente, welche auf Deci und Ryans Selbstbestimmungstheorie (1985) basieren, werden mit Umfragebögen ermittelt. Diese messen die Autonomie, die wahrgenommene Kompetenz und den Bezug zum betreffenden Thema. Insgesamt wurden 55 verschiedene Publikationen basierend auf der Selbstbestimmungstheorie gefunden.

Insgesamt zeigt sich ein ansteigendes Interesse an Motivationsforschungen zu Energiethemen. Erste Theorien beschäftigen sich mit Energieeffizienz und dessen psychologischer Erklärung. Als Beispiele seien genannt: Stern (2011), der die Korrelation zwischen Klimawandel und psychologischen Faktoren erforscht sowie Loock et al. (2013), welche die Rolle der Informationssysteme in energieeffizienten Haushalten näher untersuchen. Zusammen mit der stetig steigenden CO<sub>2</sub>-Konzentration steigt der Bedarf nach Aufklärung, mit welchem Interaktionskanal EBS am stärksten beeinflusst werden kann. Daher wird ein Erhebungsinstrument benötigt, das den Zusammenhang zwischen der Motivation zur Nutzung eines Energiesparhelferproduktes und EBS misst. Für zukünftige Forschungen im Bereich der Energieverbrauchsanzeigen im Automobil stellt sich die Forschungsfrage: **Wie muss der Prototyp eines passenden Erhebungsinstruments zur Feststellung der Motivation von Nutzern von Energieverbrauchsanzeigen im Automobil gestaltet werden?**

### 3 Erhebungsinstrument

312 von insgesamt 478 Forschungsarbeiten, die sich auf die Identifizierung von Motivation spezialisiert haben, verwendeten Fragebögen. Zusätzlich haben Fragebögen die Möglichkeit einer quantitativen Analyse mittels großer Stichprobe, Heterogenität und ortunabhängiger Einsetzbarkeit. Ein Fragebogen als Erhebungsinstrument erscheint geeignet. In der Literatur wird die 7er Likert-Skala zur Messung von Attributen präferiert und bei dieser Erstellung herangezogen.

Zuerst werden für die Erstellung des Erhebungsinstruments die zu messenden Attribute festgelegt. D.h. die Fokussierung auf das Interesse (Deci und Ryan 1985), die wahrgenommene Autonomie (Deci und Ryan 1985; Venkatesh und Davis 2000; Venkatesh et al. 2003) sowie die Kompetenz (Bandura 1977; Deci und Ryan 1985; Herzberg und Snyderman 1959; Maslow 1954; McClelland 1961; Venkatesh und Davis 2000) und die soziale Norm zur Nutzung von Energieverbrauchsanzeigen (Ajzen und Fishbein 1975; Deci und Ryan 1985; Maslow 1954; McClelland 1961; Venkatesh und Davis 2000). Des Weiteren werden die Erwartungen und die Zielsetzung zum Energiesparen abgefragt (Locke 1965, 1966, 1968; Looock et al. 2013; Mento et al. 1987; Vroom 1964; Porter und Lawler 1968). Außerdem werden die externen Einflussfaktoren auf Motivation abgefragt: das soziale Umfeld (Ajzen und Fishbein 1975; Ajzen 1991; Davis et al. 1989; Deci und Ryan 1985; Locke 1968; Venkatesh und Davis 2000; Venkatesh et al. 2003; Venkatesh und Bala 2008) und speziell dessen Reaktionen auf das Verhalten (Locke 1968; Mento et al. 1987; Skinner 1953).

Die Basis für das Erhebungsinstrument ist das *Motivation Sources Inventory (MSI)* von Barbuto und Scholl (1998), welches den motivationalen Einfluss von Belohnungen, eigenen Werten und selbstgesetzten Zielen misst. Weiterhin werden die *Selbstregulation-* sowie die *Wahrgenommene-Kompetenz-Skala*, aus der Selbstbestimmungstheorie (Deci und Ryan (1985) - SDT) stammend, mit einbezogen. Unabhängig davon ist der Sprachraum der zugrunde liegenden Literatur Englisch, in welcher die Erstellung erfolgt. Insgesamt lassen sich für den Fragebogen acht Kategorien mit 52 Fragen aus der Literatur ableiten.

Demographische Daten (1) umfassen Fragen nach dem Geschlecht, Alter und der Fahrerfahrung, welche in gefahrenen Kilometern pro Jahr und Besitz des Führerscheins in Jahren gemessen wird. Zur Ermittlung des EBSs (2) wurden sechs Fragen zu Alltagssituationen wie Wasserverbrauch, Umgang mit elektronischen Geräten und dem selbstwahrgenommenen Energieverbrauch gestellt. Die intrinsische Motivation wird in drei Bereiche aufgeteilt. Internes Selbstverständnis (MSI) (3) misst mit sechs Fragen die interne Kausalität. Dies bedeutet, wie stark eine Aktivität eigene Standards berücksichtigt. Wahrgenommene Kompetenz (SDT, vier Fragen) (4) und wahrgenommene Autonomie (SDT, zwölf Fragen) (5) messen, wie das Empfinden zur Kompetenz und Autonomie hinsichtlich der Motivation, des Wohlbefindens und der Performanz im Umgang mit Energieverbrauchsanzeigen sind. Die extrinsische Motivation wird in drei Bereiche aufgeteilt. Instrumentelle Motivation (MSI) (6) bewertet mit sechs Fragen wie Belohnungen, das soziale Umfeld und die externe Rückmeldung von der eigenen Person erfasst werden. Externes Selbstverständnis (MSI) (7) misst mit sechs Fragen den Grad des externen Einflusses von Bezugspersonen wie Eltern oder Vorbildern. Internalisierung von Zielen (MSI) (8) misst mit sechs Fragen, in wie weit die individuellen Ziele mit der eigenen Wertorientierung konsistent sind.

Die Validierung erfolgt unter der Annahme, dass wenn eine Person in der Kategorie (2) ein hohes Maß an EBS hat, dann ist diese entweder intrinsisch, extrinsisch, oder intrinsisch und extrinsisch motiviert, um Energie zu sparen. Dies bedeutet, dass die Energieverbrauchsanzeige im Auto austauschbar wäre und somit der Fragebogen verallgemeinert werden kann. Dazu werden die

Ergebnisse der intrinsischen (3,4,5) und extrinsischen Kategorien (6,7,8) zusammengefasst und jeweils der Mittelwert gebildet. Danach wird die Korrelation zur Kategorie (2) EBS geprüft.

Am seit dem 26. August 2015 stattfindenden Pretest haben 49 Personen teilgenommen. Der Fragebogen hat ein durchschnittliches Cronbachs Alpha von 0,74. Die Teilnehmer sind Studierende zwischen 20 und 30 Jahren, davon 20 Frauen und 29 Männer. Zwischenergebnisse zeigen, dass es ein generelles Interesse an Energieeffizienz gibt (31 von 49 Personen gaben einen Wert von 5 oder höher auf der 7er Likert-Skala an). Je höher die gefahrenen Kilometer desto höher ist die Motivation zur Benutzung einer Energieverbrauchsanzeige im Automobil, um Energie zu sparen (Pearson Korrelation von 0,296). Dies bedeutet, dass das EBS mit gefahrenen Kilometern steigt. Es gibt bisher eine signifikante Korrelation zwischen dem EBS und dem internen Selbstverständnis (Pearson Korrelation 0,317). Dies bedeutet, dass je energiesparender agiert wird, desto ausgeprägter ist das EBS der Person. Dieser Effekt scheint sich durch die gesetzten Normen und sozialen Standards beeinflussen zu lassen. Es ist nicht klar, welche extrinsischen Motivatoren z.B. Feedback via Interaktionskanäle zu einem positiven oder negativen Effekt führen.

#### **4 Diskussion**

Es wird vorausgesetzt, das gilt, dass eine Person mit einer hohen Motivation zum Energiesparen auch ein hohes EBS besitzt. Der Fragebogen hat insgesamt 52 Objekte. Dieses kann eine hohe Durchführungszeit implizieren und damit zu einer hohen Abbruchquote führen. Was die Ergebnisse verfälschen könnte und in einem weiteren Schritt zu prüfen ist. Ferner könnte der Fragebogen in ins Deutsche übersetzt und validiert werden, um mögliche Verfälschungen ausschließen zu können.

Eine erste Evaluierung des Erhebungsinstruments zeigt, dass ein verbessertes EBS möglich ist. Motivationsforschung bietet Grund zur Annahme, dass maßgeblich intrinsische Ideale und Wertvorstellungen die entscheidenden Intentionen zum Energiesparen sind. Da das EBS mit den gefahrenen Kilometern steigt, ist es notwendig herauszufinden, welche Informationskanäle genutzt werden können, um das EBS positiv zu beeinflussen. In weiteren Schritten können Experimente genutzt werden um die interne Validität zu erhöhen und geeignete Verbrauchsanzeigen identifiziert werden. Unabhängig davon ermöglicht es Typen von EBS (Kunden) zu identifizieren und somit kundenorientierte Energieverbrauchsanzeigen zu gestalten.

In einem zukünftigen Schritt ist das Erhebungsinstrument zu validieren. Die Stichprobengröße ist bisher sehr gering und ist für eine bessere Aussagekraft zu erhöhen. Außerdem sind altersbedingte Unterschiede in einem weiteren Schritt zu berücksichtigen.

Neben der Notwendigkeit im Automobilbereich könnte in einem zukünftigen Schritt eine verfeinerte, allgemeingültige Version des Erhebungsinstruments in allen energiesensitiven Bereichen eingesetzt werden, um eine Korrelation zwischen Energiesparprodukten und deren Einfluss auf Nutzungsmotivation zu erhalten. Zunächst dient der Fragebogen jedoch in einem Laborexperiment.

#### **5 Literatur**

- Adams JS (1963) Toward an Understanding of Inequity. *Journal of Abnormal and Social Psychology* 67 (5):14
- Ajzen I (1991) The Theory of Planned Behavior. *Organizational Behavior and Human Decision Processes* 50:179-211

- Ajzen I, Fishbein M (1975) *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*. Addison-Wesley, Reading, MA
- Alderfer CP (1969) An Empirical Test of a New Theory of Human Needs. *Organizational Behavior and Human Performance* 4 (1):142-145
- Bandura A (1977) Self-Efficacy: Toward an Unifying Theory of Behavioral Change. *Psychological Review* 84 (2):191-215
- Bandura A (1999) Social Cognitive Theory: An Agentic Perspective. *Asian Journal of Social Psychology* 2:21-41
- Barbuto JE, Scholl RW (1998) Motivation Sources Inventory: Development and Validation of New Scales to Measure an Integrative Taxonomy of Motivation. *Psychological Reports* 82:1011-1022
- Becker LJ (1978) Joint Effect of Feedback and Goal Setting on Performance: A Field Study of Residential Energy Conservation. *Journal of Applied Psychology* 63 (4):428-433
- Brocke J, Simons A (2009) Reconstructing The Giant: On the Importance of Rigour in Documenting the Literature Search Process. *ECIS 2009*:1-14
- Davis F, Bagozzi RP, Warshaw PR (1989) User Acceptance Computer Technology - A Comparison of Two Theoretical Models. *Management Science* 35 (8):982-1003
- Deci EL (1971) Effects of Externally Mediated Rewards on Intrinsic Motivation. *Journal of Personality and Social Psychology* 18:105-115
- Deci EL (1972a) The Effect of Contingent and Non-Contingent Rewards and Controls on Intrinsic Motivation. *Organizational Behavior and Human Performance* 8:217-229
- Deci EL (1972b) Intrinsic Motivation and Extrinsic Reinforcement, and Inequity. *Journal of Abnormal and Social Psychology* 22:113-120
- Deci EL, Ryan RM (1985) *Intrinsic Motivation and Self-Determination in Human Behavior*. Springer Science & Business Media,
- Elliot AJ, Covington MV (2001) Approach and Avoidance Motivation. *Educational Psychology Review* 13 (2):73-92
- Harlow HF (1953) Learning by Rhesus Monkeys on the Basis of Manipulation-Exploration Motives. *Science* 117:466-467
- Hebb DO (1955) Drives and the C. N. S. (Central Nervous System). *Psychological Review* 62:243-254
- Herzberg F, Snyderman BB (1959) *The Motivation to Work*. Wiley, New York
- Hull C (1943) *Principles of Behavior*. Appleton-Century-Crofts, New York
- IPCC (2007) *Climate Change 2007: Synthesis Report*. Intergovernmental Panel on Climate Change, Geneva, Switzerland
- James W (1890) *Principles of Psychology*. H. Holt and Company, New York
- Levesque C, Copeland KJ, Pattie MD, Deci EL (2010) Intrinsic and Extrinsic Motivation. *International Encyclopedia of Education* 3 (2010):618-623



- Locke EA (1965) A Test of Atkinson's Formula for Predicting Choice Behavior. *Psychological Reports* 16:963-964
- Locke EA (1966) The Relationship of Intentions To Level of Performance. *Journal of Applied Psychology* 50 (1):60-66
- Locke EA (1968) Toward a Theory of Task Motivation and Incentives. *Organizational Behavior and Human Performance* 3 (2):157-189
- Loock C-M, Staake T, Thiesse F (2013) Motivating Energy-Efficient Behavior With Green IS: An Investigation of Goal Setting And The Role of Defaults. *MIS Quarterly* 37 (4):1313-1332
- Maslow AH (1943) A Theory of Human Motivation. *Psychological Review* 50:370-396
- Maslow AH (1954) *Motivation and Personality*. Harper & Row, New York
- McClelland DC (1961) *The Achieving Society*. D. Van Nostrand, Princeton, NJ
- McDougall W (1908) *An Introduction to Social Psychology*. John W. Luce & Co., Boston
- Mento A, Steel R, Karren R (1987) Meta-Analytic Study of the Effect of Goal Setting on Task Performance: 1966-1984. *Organizational Behavior and Human Decision Processes* 39 (1):52-83
- National-Research-Council (2010) *Advancing the Science of Climate Change*. National Academies Press, Washington, D.C.
- Porter LW, Lawler EE (1968) *Managerial Attitudes and Performance*. Richard D. Irwin, Inc., Homewood, IL
- Seligman C, Darley JM (1977) Feedback as a Means of Decreasing Residential Energy Consumption. *Journal of Applied Psychology* 62 (4):363-368
- Skinner BF (1938) *The Behavior of Organisms: An Experimental Analysis*. Prentice-Hall, Englewood Cliffs, NJ
- Skinner BF (1953) *Science and Human Behavior*. Macmillan, New York
- Stern PC (2011) Contributions of Psychology to Limiting Climate Change. *American Psychologist* 66 (4):303-331
- Venkatesh V, Davis F (2000) A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science* 46 (2):186-204
- Venkatesh V, Morris MG, Davis GB, Davis F (2003) User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly* 27 (3):425-478
- Vroom VH (1964) *Work and Motivation*. Wiley, New York, NY
- Watson J (1913) *Psychology as the Behaviorist Views it*. Columbia University
- Woodworth RS (1917) *Dynamic Psychology*. Columbia University - The Jesup Lectures
- WMO (2014) Record Greenhouse Gas Levels Impact Atmosphere and Oceans. Press Release. [https://www.wmo.int/pages/mediacentre/press\\_releases/pr\\_1002\\_en.html](https://www.wmo.int/pages/mediacentre/press_releases/pr_1002_en.html).
- Zuckerman M (1979) *Sensation Seeking: Beyond the Optimal Level of Arousal*. Hillsdale, NJ, Erlbaum



# The Three Kinds of Trust in the Sharing Economy

Florian Hawlitschek<sup>1</sup>, Christian Peukert<sup>1</sup>, Timm Teubner<sup>1</sup>, and Christof Weinhardt<sup>1</sup>

<sup>1</sup> Karlsruhe Institute of Technology, Institute of Information Systems and Marketing (IISM),  
{florian.hawlitschek; timm.teubner; weinhardt}@kit.edu, christian.peukert@me.com

## Abstract

Trust has been in the focus of research on B2C e-commerce in the last decade. Nowadays the rise of consumer-to-consumer (C2C) markets in the context of the sharing economy poses new challenges and questions. This work outlines a conceptual research model for the role of trust regarding consuming and providing intentions in such markets. Our model differentiates between three substantial kinds of trust, that is, towards *peer*, *platform*, and *product* (3P). Moreover, we consider and differentiate between both the consumer and provider perspective. A first exploratory evaluation of trust-related design patterns on three platforms in the sharing economy landscape (Airbnb, Tamyca, eBay) provides insights on how the «3P» are currently addressed in the field.

## 1 Introduction

*“Sharing, whether with our parents, children, siblings, life partners, friends, coworkers, or neighbors, goes hand in hand with trust and bonding.”* (Belk 2010, 717)

While sharing is almost as old as mankind (Sahlins 1972) the sharing economy, intermediated by Internet and mobile technology, is a phenomenon of the 21<sup>st</sup> century. In fact, driven by the facilitating role of peer-to-peer platforms and Information Systems (IS), its rise is changing the consumption behavior of millions of people around the globe. Large sharing economy platforms such as Airbnb ([www.airbnb.com](http://www.airbnb.com)) exceed their figures every year. Research, however, is struggling to keep up with this rapid development. Even the term *sharing economy* itself still lacks a widely accepted and precise definition. In the IS community it is primarily used as an umbrella term for phenomena such as *Collaborative Consumption* (Botsman and Rogers 2010), *Commercial Sharing Systems* (Lamberton and Rose 2012), or *Access-Based Consumption* (Bardhi and Eckhardt 2012).

Sharing is closely related to trust (Belk 2010). In the context of the sharing economy however, trust is assumed to play a crucial role and was even referred to as its currency (Botsman and Rogers 2010). Large international business consultancies also agree on that fact: “To share is to trust. That, in a nutshell, is the fundamental principle [...]” stated *Roland Berger* (in the *Think Act Shared Mobility*, July 2014). One year later *PwC* stated that “[...] convenience and cost-savings are beacons, but what ultimately keeps this economy spinning—and growing—is trust.” (in the *Consumer Intelligence Series: The Sharing Economy*, April 2015). After several years of fundamental research regarding trust in business to consumer e-commerce (e.g., Gefen 2000,

McKnight and Chervany 2002, Gefen and Straub 2004), an increasing number of scholars has started to explore the role of trust in consumer-to-consumer (C2C) e-commerce (e.g., Jones and Leonard 2008, Lu et al. 2010, Leonard 2012). It is one, if not *the* important driving factor for the long term success of C2C platforms (Strader and Ramaswami 2002). However, trust is a multifaceted and complex construct—often hard to pin down (Keen et al. 1999). While in traditional (B2C) e-commerce it can be understood as a willingness to depend on an online vendor from an IS perspective (Gefen and Straub 2004), the picture is more complex for C2C markets. Internet users engage in interactions with more than one party. The vendor’s role is taken by another private individual, renting out his or her car, apartment, or other equipment. The platform, however, acts as a broker and mediator between both sides in the market, and may appear trustworthy or not. Moreover, even the product itself (think for example of a privately rented apartment or car) may be subject to trust concerns. This paper thus outlines a conceptual research model for the role of trust in C2C markets. Our model differentiates between two perspectives (consumer and provider), as well as three kinds of trust (in peer, platform, and product). An exploratory evaluation of trust-related design patterns on three platform providers in the sharing economy landscape illustrates how the suggested «3P» are currently addressed in the field. We contribute to research on trust in online environments by shedding light on the kinds and dimensionality of trust in the sharing economy.

The remainder of this paper is structured as follows. Section 2 presents the theoretical background of trust in C2C markets, building on IS theories of trust in traditional e-commerce context. We then present our model and derive its central hypotheses. In Section 3, we present the results of our exploratory evaluation of trust-related design patterns. We summarize our findings, and draw a brief conclusion in Section 4.

## 2 Theoretical Background & Research Model

Linking social presence to consumer trust, Gefen and Straub (2004) made a significant contribution in the area of trust in B2C e-commerce that was frequently cited and used as foundation for succeeding research models and approaches. Their model focusses on human behavior in the context of “traditional” electronic commerce, i.e., an Internet user facing the website of an e-vendor. Trust in this context is introduced as a multidimensional construct which differentiates between the four dimensions *ability (ABLY)*, *integrity (INTG)*, *benevolence (BNVL)*, and *predictability*. However, caused by the relationship of the parties concerned in a transaction, other aspects are focused on in studies dealing with trust in C2C e-commerce. Lu et al. (2010) analyzed how trust affects purchase intentions in the context of C2C buying in virtual communities. They found that especially the community members’ trustworthiness influenced purchase intentions. For this purpose, their research model differentiates between the constructs *trust in members* and *trust in website/vendor* of the virtual community. Both constructs were separated into three dimensions: *ABLY*, *INTG*, and *BNVL*. For the construct *trust in members* *INTG* and *BNVL* were merged into a single dimension. Jones and Leonard (2008) in contrast considered *C2C trust* as a single, one-dimensional construct and hypothesized internal (*natural propensity to trust*, *perception of website quality*) and external (*other’s trust*, *third party recognition*) as influencing factors within C2C e-commerce settings. In a more recent study, Leonard (2012) distinguishes between the two one-dimensional constructs *trust in seller* and *trust in buyer* which, along with *risk* of both, seller and buyer are hypothesized to influence selling or buying attitudes. However, as depicted in Table 1, none of the above mentioned models covers all kinds, dimensions, and perspectives that appear as

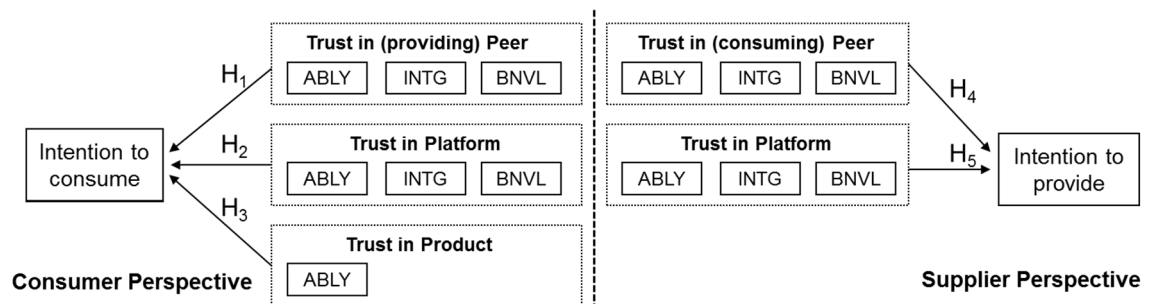
relevant in the context of the sharing economy. Hence, we suggest a comprehensive conceptual research model of trust in C2C sharing economy platforms.

	KINDS OF TRUST			PERSPECTIVES	
	<i>peer</i>	<i>platform</i>	<i>product</i>	consumer	provider
Gefen and Straub, 2004		x		x	
Jones and Leonard, 2008	x			x*	x*
Lu et al., 2010	x	x		x*	x*
Leonard 2012	x			x	x
This work	x	x	x	x	x

**Table 1: Literature on kinds and perspectives for trust in the sharing economy (\*joint perspective)**

### A Research Model of Trust in C2C Sharing Economy Platforms

Based on the above, we propose the conceptual research model as depicted in Figure 1. Our key objective is to describe how trust influences users' intentions to transact on sharing economy platforms. To this end, we differentiate the perspectives of consumers and suppliers. Moreover, the model distinguishes between three different kinds of trust: towards *Peer*, *Platform*, and *Product* («3P») —represented by the established dimensions *ABLY*, *INTG*, and *BNVL*, respectively. Within the scope of this work, we present a conceptual research model as a simplified basis for future research. Further aspects such as trust transfer and antecedents of trust (Lu et al. 2010) will be addressed in future work.



**Figure 1: Research model for trust in C2C markets (ABLY: ability; INTG: integrity; BNVL: benevolence)**

#### Consumer perspective

*Trust in (providing) peer* describes whether the provider has the skills and competences to execute his part of the transaction, and whether he is considered as a transaction partner of high integrity and benevolence (Pavlou and Mendel 2006). The intention to complete a transaction was found to depend on trust in the offering peer (Lu et al. 2010). The constructs *INTG* («the provider keeps his word») and *BNVL* («the provider is interested in satisfying the customer») are closely related as a benevolent provider will most likely also exhibit high levels of integrity and vice versa. Several scholars have thus employed joint constructs to assess the general notion, e.g., in the context of virtual communities (Ridings et al. 2002, Lu et al. 2010). The general notions of integrity and benevolence are particularly important in C2C markets—compared to B2C—for at least two interacting reasons. First, the supplying peer will most likely not appear as a legal entity but as a private person. In many countries regulative buyer protection does not exist or is limited for private-to-private transactions. Second, customers in today's C2C market interactions are often put into a particular vulnerable position, where—e.g. in the context of apartment and ride sharing—they strongly depend on the desirable behavior and task fulfillment of the supplying peer: Who wants to end up in a foreign city late at night, discovering that the booked and paid apartment simply does

not exist or that the driver does not show up? Another important aspect is ability. Given that a transaction partner is well-meaning, it could still be that he or she is simply lacking the skills to properly (or safely) complete the task—think for example of amateur or hazardous UBER ([www.uber.com](http://www.uber.com)) drivers who might unintentionally endanger a customer's safety. This speaks in favor of the conjecture that trust (based on ability, integrity, and benevolence) towards the supplying peer positively affects a user's intention to consume in a C2C market. We hence hypothesize that:

**H<sub>1</sub>:** *Trust in the (supplying) peer positively affects intention to consume.*

According to Gefen (2002), *trust in platform* is also based on beliefs about ability, integrity, and benevolence of a website or vendor. In contrast to B2C the platform operator in C2C markets primarily acts as a mediator between the peers. Ability here could refer to whether the platform successfully finds and connects transaction partners, i.e., its adoption. Secure and reliable data handling is another important aspect. Perceptions of a platform's integrity and benevolence, in turn, could be linked to how much it charges its users, the design of user support, excessive email spamming, third party access to user data, and its general reputation, for instance, for being a "data kraken" or exploiting providers. To find a suitable offer, a user typically creates an account (providing private data such as name, credit card information, email, etc.). Privacy calculus theory states the privacy risk involved with this behavior is weighted against its benefits, where trusting beliefs towards the platform operator are positively associated with intention to disclose (Krasnova et al. 2012; Dinev and Hart 2006). Moreover, Gefen (2002) found that trust in platform's ability positively affects window-shopping intentions of consumers and that trust in the integrity as well as benevolence affects the purchase intention. We hence suggest that:

**H<sub>2</sub>:** *Trust in the platform positively affects intention to consume.*

*Trust in product* describes how the product itself is perceived as reliable by the (potential) consumer. Comer et al. (1999) defined "product trust [as] the belief that the product/ service will fulfill its functions as understood by the buyer" (p. 62). We transfer this notion to C2C platforms. A rented car needs to work for obvious reasons of convenience and safety, a rented or purchased good is expected to fulfill its purpose, and also a rented apartment needs to be functional. Since the product is an inanimate object, it does not have a will or intention. Its functionality and quality are covered by the trust dimension of ability. Our third hypothesis hence states:

**H<sub>3</sub>:** *Trust in the product positively affects intention to consumer.*

### **Supplier perspective**

As most C2C platforms work on the basis of mutual agreement to trigger a transaction, also the provider's *trust in the consuming peer* is of importance. This becomes particularly evident for peer-to-peer rental services as the provider cedes her car, apartment, or other resource (the platform *Rover.com* even connects dog owners and sitters) to another person for use and has no effective control over it for the agreed period of time. Consequently, entrusting personal belongings—one's home, car, let alone a pet—to an unknown stranger requires that the supplier trusts in the ability of the consumer: On the one hand, being convinced by the skills and on the other hand by the knowledge the consumer owns (Lu et al. 2010). Nevertheless, without the supplier trust in the integrity and benevolence of the consuming peer, an agreement is hard to achieve. Against the background of the two constructs *INTG* (‹the consumer keeps his word›) and *BNVL* (‹the consumer keeps the suppliers interests in mind›) this means that the supplier would need to be convinced that her possessions are neither used for purposes that were not agreed nor over- or abused. Think for example of renting out your car at Tamyca ([www.tamyca.de](http://www.tamyca.de)) to someone who owns a driver's

license—which technically means the person is able to drive a car—but conveys the impression that he or she does neither care about the exact time of returning, nor about the condition of the car. We therefore suggest:

**H<sub>4</sub>:** *Trust in the (consuming) peer positively affects intention to provide.*

In accordance with the train of thought leading to the three dimensions of trust from the consumer perspective, provider’s *Trust in the platform* also rests upon the constructs *ABLY*, *INTG*, and *BNVL*. The platform’s ability in this context can be understood as a competence or qualification for seamless communication and service operation, i.e. the successful mediation between peers. Suppliers might for example expect an adequate pre-selection of requests by the platform operator as well as a functional and easy-to-use booking, payment, and reputation system. Aspects, such as reliability (especially regarding data privacy and potential claims) or safeguarding of supplier interests (e.g. legal certainty and payments) are reflected in the integrity and benevolence dimension. From a supplier’s perspective mechanisms to absorb risks of resource damage, exemplarily by a standardized insurance coverage and transparent profit-sharing mechanisms might increase the trust in a certain platform. Furthermore, communication protocols facilitating a supplier’s data security so that privacy is not threatened unduly also appear beneficial in terms of promoting trust towards a platform. Eventually, we suggest that *trust in platform* also plays a role for the supplier’s intention to commit a transaction:

**H<sub>5</sub>:** *Trust in the platform positively affects intention to provide.*

As the offered product belongs to the supplying peer, its abilities should be known by the supplier. Therefore, a trust dimension from the supplier’s point of view is not considered as relevant. Whenever possible, we reused well-established constructs in our survey development, whereas we adopted the items to fit the context of the sharing economy and C2C platform. A comprehensive list of all items and constructs can be requested from the authors.

### 3 Evaluation of Trust-Related Design Patterns

In the following, we consider the use of trust-related design patterns on the platforms Airbnb, Tamyca, and eBay’s local redistribution market ([www.ebay-kleinanzeigen.de](http://www.ebay-kleinanzeigen.de)) as representatives of a much wider range of different sharing economy platforms (c.f. Botsman 2010). In the literature, design patterns are described as generic solutions for commonly occurring problems (Granlund et al. 2001). Trust-relevant design patterns were exemplarily extracted and assigned to one of the «3P» (trust in *platform*, *peer*, or *product*) following an explorative website click path through the above mentioned websites. Some of the patterns are associated with more than one type of trust. All data was collected from the respective websites in January 2015. An overview of the identified patterns and their application on Airbnb, Tamyca, and eBay is provided in Table 2 and Table 3.

	PIC	NAM	DES	VRF	MMB	RTN	TME	REV	EVA	AWR	CRT	MES	PRO
Airbnb	x	x	x	x	x	x	x	x	x	x		x	x
Tamyca	x	x	x	x		x	x	x			x	x	
eBay			x		x							x	x

**Table 2: Design patterns for trust in peer**

Tamyca and Airbnb use similar design patterns to promote trust among *peers*. Some of these are used symmetrically both for the supply and demand side. Both sides present themselves using a profile picture (PIC), their actual name (NAM), and a textual self-description (DES). Moreover,

both platforms verify their users’ IDs (VRF). Tamyca also asks potential drivers to submit a copy of their driver’s license. The platforms differ with regard to minor information bits provided on the websites, as for instance, the user’s duration of membership (MMB). The supplying peer’s trustworthiness is assessed by star-rating systems (RTN) and reports on the average response time to customer requests (TME). Furthermore, interaction partners can write reviews based on their experiences (REV). Airbnb specifically points out the number of evaluations (EVA) and particularly well rated suppliers receive an award (AWR) in form of a “superhost” badge which is prominently shown on their profile. Tamyca demands their users to meet special criteria (CRT), including a minimum time in possession of a driver’s license, similarly to car rental companies. It is quite salient that eBay almost completely refrains from using typical peer trust elements—except from the opportunity to contact the supplier via a messaging system (MES) and the provision of information on their account’s history. Furthermore, eBay explicitly displays the number of products currently online (PRO) which might shed light on the supplier’s experience in providing products.

	INS	SPP	STP	GUI	FRD		PIC	DES	REV	RTT	VIS
Airbnb	x	x	x	x		Airbnb	x	x	x		
Tamyca	x	x	x	x		Tamyca	x	x	x		
eBay				x	x	eBay	x	x		x	x

**Table 3: Design patterns for trust in platform (left) and trust in product (right)**

For trust in *platform*, Tamyca and Airbnb again use similar trust-related design patterns. Both providers cooperate with an insurance company (INS) and offer professional support (SPP) to assure some degree of certainty to their users. Additionally, they offer a detailed step-by-step description (STP) of the user interaction process. In contrast, eBay does not use these elements. Here, products are sold, not rented, a usage insurance simply does not apply. All platforms list a few things the buyer and sometimes the supplier should pay attention to, and thereby guide their actions (GUI). A button to report fraudulent use or products (FRD), is explicitly offered by eBay. However, Airbnb’s and Tamyca’s users can always report misuse by e-mail or contact forms. Regardless of the platform, similar elements are used to describe the *product*, including pictures (PIC) and descriptions (DES). Again, product reviews (REV) do not apply for eBay as its interactions are one-shot. Instead, the platform provides two additional metrics: A product’s online retention time (RTT) and its number of visitors (VIS).

## 4 Conclusion and Outlook

Within the scope of this research-in-progress paper, we proposed a research model for the role of trust in C2C sharing economy platforms that is based on the «3P» of trust, i.e., towards *platform*, *peer*, or *product*—represented by the dimensions *ABLY*, *INTG*, and *BNVL*. It incorporates both, the consumers’ and suppliers’ intentions to consume or provide a resource, as both are represented by private, i.e. non-professional, persons. In a first exploratory evaluation of trust-related design patterns on the platforms Airbnb, Tamyca, and eBay Kleinanzeigen, we illustrate how platform providers address the three mentioned kinds of trust for both roles in the market. In future research, we plan to develop and validate a survey-based measurement model for the suggested trust constructs and to enhance our conceptual research model by integrating aspects such as trust transfer and antecedents of trust (Lu et al. 2010). From an IS perspective the role of trust in *platform*, *peer*,



or *product* on different C2C platforms in the context of the sharing economy can subsequently be evaluated in order to develop guidance regarding the use of appropriate trust-related design patterns.

## 5 References

- Bardhi F, Eckhardt GM (2012) Access-based consumption: the case of car sharing. *Journal of consumer research* 39(4): 881–898
- Belk R (2010) Sharing. *Journal of consumer research* 36(5): 715–734
- Botsman R, Rogers R (2010) What's mine is yours. The rise of collaborative consumption. Collins
- Comer JM, Plank RE, Reid DA, Pullins EB (1999) Methods in sales research: perceived trust in business-to-business sales: a new measure. *JPSSM* 19(3): 61–71
- Dinev T, Hart P (2006) An extended privacy calculus model for e-commerce transactions. *Information systems research* 17(1): 61–80
- Gefen D (2000) E-commerce: the role of familiarity and trust. *Omega* 28(6): 725–737
- Gefen D (2002) Customer loyalty in e-commerce. *Journal of the association for information systems* 3(1): 27–51
- Gefen D, Straub DW (2004) Consumer trust in B2C e-commerce and the importance of social presence: experiments in e-products and e-services. *Omega* 32(6): 407–424
- Granlund A, Lafrenière D, Carr DA (2001) A pattern-supported approach to the user interface design process. In: *Proceedings of HCI international, New Orleans*
- Jones K, Leonard LN (2008) Trust in consumer-to-consumer electronic commerce. *Information & management* 45(2): 88–95
- Keen P, Ballance G, Chan S, Schrupp S (1999) *Electronic commerce relationships: trust by design*. Prentice Hall PTR, Upper Saddle River
- Krasnova H, Veltri NF, Günther O (2012) Self-disclosure and privacy calculus on social networking sites: the role of culture. *Business & information systems engineering* 4(3): 127–135
- Lamberton CP, Rose RL (2012) When is ours better than mine? A framework for understanding and altering participation in commercial sharing systems. *Journal of marketing* 76(4): 109–125.
- Leonard, LN (2012) Attitude influencers in C2C e-commerce: buying and selling. *Journal of computer information systems* 52(3): 11–17
- Lu Y, Zhao L, Wang B (2010) From virtual community members to C2C e-commerce buyers: trust in virtual communities and its effect on consumers' purchase intention. *Electronic commerce research and applications* 9(4): 346–360
- Mcknight DH, Chervany NL (2002) What trust means in e-commerce customer relationships: an interdisciplinary conceptual typology. *International journal of electronic commerce* 6(2): 35–59
- Pavlou PA, Fygenson M (2006) Understanding and predicting electronic commerce adoption: an extension of the theory of planned behavior. *MIS quarterly* 30(1): 115–143
- Ridings CM, Gefen D, Arinze B (2002) Some antecedents and effects of trust in virtual communities. *The Journal of strategic information systems* 11(3): 271–295
- Sahlins MD (1972). *Stone Age Economics*. Aldine, New York
- Strader TJ, Ramaswami SN (2002) The value of seller trustworthiness in C2C online markets. *Communications of the ACM* 45(12): 45–49.



# **Der ISO27-Sicherheitsprozess: Ein Referenzprozess zur Umsetzung der ISO/IEC 27001**

**Marlen Hofmann<sup>1</sup> und Andreas Hofmann<sup>2</sup>**

<sup>1</sup> Martin-Luther-Universität Halle-Wittenberg, marlen.hofmann@gmail.com

<sup>2</sup> CISO27, Leipzig, info@ciso27.de

## **Abstract**

Neue regulatorische Anforderungen und die gestiegene Bedeutung von Informationssicherheit im Allgemeinen motivieren viele Unternehmen zur Einführung eines Informationssicherheitsmanagementsystems (ISMS) gemäß den Anforderungen der ISO/IEC 27001 (ISO 27001). Die Norm enthält jedoch keine konkreten Ausgestaltungshinweise für einen durchgängigen Sicherheitsprozess, sodass ISMS-Praktiker häufig auf die IT-Grundschutz Vorgehensweise des Bundesamtes für Sicherheit in der Informationstechnik (BSI) zurückgreifen. Hierdurch können jedoch methodische Brüche und Inkonsistenzen im ISMS-Vorgehen entstehen und einzelne Anforderungen der Norm nicht oder nur unzureichend berücksichtigt werden. Die Zielsetzung der hier vorgestellt Forschungsarbeit ist es daher, einen Referenzprozess für einen normkonformen ISO27-Sicherheitsprozess als Pendant zum BSI-Sicherheitsprozess zu entwickeln, der "by design" die Anforderungen der ISO 27001 erfüllt.

## **1 Executive Summary**

Neue regulatorische Anforderungen (insb. das IT-Sicherheitsgesetz und der IT-Sicherheitskatalog) aber auch die gestiegene Bedeutung von Informationssicherheit im Allgemeinen motivieren Unternehmen nahezu aller Branchen und Größen zur Einführung und Zertifizierung eines Informationssicherheitsmanagementsystems (ISMS). Zielsetzung des ISMS ist es, unternehmenskritische Informationen vor Verlust, Manipulation oder Kenntnisnahme durch unberechtigte Dritte zu schützen. Dazu werden ein Vielzahl von Handlungsanweisungen mit Bezug zur Informationssicherheit erstellt, unterschiedlichste ISMS-relevante Analysen durchgeführt und Maßnahmen zum Schutz von bedrohten Informationen umgesetzt (Vgl. Kersten et al. 2013, Klipper 2015).

Anforderungen zur Ausgestaltung des ISMS finden sich insbesondere in den IT-Grundschutz-Standards (GS-Standards) des Bundesamtes für Sicherheit in der Informationstechnik (BSI) und der internationalen DIN-Norm ISO/IEC 27001 (ISO 27001). Obwohl sich beide ähneln, weisen sie dennoch wesentliche Unterschiede auf. Die ISO 27001 fokussiert auf einen angemessenen und risikoorientierten Schutz der Informationen vor Informationssicherheitsrisiken (IS-Risiken) und enthält sehr generische und allgemein formulierte Anforderungen an ein ISMS (Vgl. ISO/IEC 27001, 2015). In den GS-Standards werden die ISMS-relevanten Aktivitäten genauer spezifiziert

und in einem Sicherheitsprozess beschrieben. Darüber hinaus schreiben die GS-Standards die Umsetzung von konkreten Schutzmaßnahmen vor, die zu einem "normalen bis hohen" Standardschutzniveau führen sollen (Vgl. BSI 2015).

Häufig werden die GS-Standards jedoch als zu aufwendig und inflexibel in der Umsetzung diskutiert, sodass sich die ISO 27001 zunehmender Beliebtheit erfreut. Die Unternehmen schätzen insbesondere die Freiheitsgrade und die Möglichkeiten zur individuellen Ausgestaltung eines, auf die jeweilige Unternehmenssituation zugeschnittenen ISMS (Vgl. Lange 2015, Reppner 2010). Da es der Norm jedoch an konkreten Vorgaben zur praktischen Ausgestaltung mangelt, werden nicht selten die GS-Standards hinzugezogen, um konkrete Arbeitsschritte, Arbeitsschrittfolgen und Methoden aus dem Sicherheitsprozess des BSI abzuleiten (Vgl. Kersten et al. 2013). In der Praxis solcher "ISMS-Mischformen" kommt es jedoch häufig zu methodischen Brüchen und Inkonsistenzen im ISMS-Vorgehen, da sich die Elemente des BSI-Sicherheitsprozess nicht ohne Weiteres in ein ISO 27001-konformes Vorgehen überführen lassen.

Zielsetzung der Forschungsarbeit ist es daher, mit dem ISO27-Sicherheitsprozess einen normkonformen, konsistenten Referenzprozess zu entwickeln, der "by design" die Anforderungen der ISO 27001 erfüllt. Gemäß gestaltungsorientiertem Forschungsparadigma findet sich in diesem Beitrag ein erster Entwurf des ISO27-Sicherheitsprozesses, der mit der wissenschaftlichen Community diskutiert, verfeinert und in den nächsten Forschungsiterationen gemeinsam mit Praxispartnern evaluiert soll (Vgl. Österle et al. 2010). Der Beitrag ist wie folgt gegliedert: Im folgenden Kapitel werden die Grundlagen zum Informationssicherheitsmanagement dargestellt und anschließend der Entwurf des ISO27-Sicherheitsprozesses präsentiert. Im vierten Kapitel findet sich eine kurze Zusammenfassung und der Forschungsausblick.

## **2 Informationssicherheit & Informationssicherheitsmanagement**

Informationssicherheit beschreibt einen Zustand, der die Verfügbarkeit, Integrität und die Vertraulichkeit von Informationen (Schutzziele) zusichert. Dabei ist unerheblich, ob die Informationen elektronisch verarbeitet, auf Papier abgedruckt oder in den Köpfen einzelner Menschen gespeichert sind. Ziel des Informationssicherheitsmanagements ist es entsprechend, den Zustand der Informationssicherheit durch geeignete Maßnahmen herzustellen und aufrechtzuerhalten. Das bedeutet konkret, dass die Informationen vor Verlust (Verfügbarkeit), Manipulation (Integrität) und vor Kenntnisnahme durch unberechtigte Dritte (Vertraulichkeit) geschützt werden. Im Rahmen des ISMS werden die hierfür erforderlichen Aktivitäten und Maßnahmen geplant, initiiert, koordiniert und deren Wirksamkeit überprüft (Vgl. Kersten et al. 2013, Klipper 2015, BSI 2015).

Anforderungen an den Aufbau und Betrieb eines ISMS werden insbesondere in der ISO 27001 und in den GS-Standards 100-1, 100-2 und 100-3 beschrieben. Die ISO 27001 ist eine international verbreitete Norm, die aus dem, im Jahr 1998 erstmals veröffentlichten British Standard 7799-2 hervor ging und deren Fokus auf Informationssicherheit im Allgemeinen liegt. Die Norm ist Bestandteil der ISO 27000 Normenreihe und beschreibt auf allgemeiner und generischer Ebene, welche Anforderungen erfüllt werden müssen, um ein normkonformes ISMS zu betreiben. Auf konkrete Umsetzungshinweise zur Ausgestaltung der notwendigen ISMS-Arbeitsschritte sowie auf eine prozessuale Beschreibung dieser wird in der Norm verzichtet (Vgl. ISO/IEC 27001, 2015, Kersten et al. 2013).

Im Vergleich dazu haben die GS-Standards deutsche Wurzeln und entstammen dem Anfang der 90er Jahre vom BSI veröffentlichten IT-Grundschutzhandbuch, dessen Fokus insbesondere auf IT-Sicherheit und Datenschutz lag. Die Anforderungen an den Aufbau und Betrieb eines ISMS nach BSI werden in den GS-Standards 100-1, 100-2 und 100-3 beschrieben. Die im Vergleich zur ISO 27001 sehr umfangreiche ISMS-Kerndokumentation enthält konkrete methodische Vorgaben und Umsetzungshinweise sowie eine prozessuale Beschreibung des BSI-Sicherheitsprozess, in dem die einzelnen ISMS-Aktivitäten in eine Reihenfolge gebracht werden (Vgl. BSI 2015, Kersten et al. 2013).

Die strengen Anforderungen des BSI sowie der Umfang der GS-Standards werden unter ISMS-Praktikern häufig als "zu komplex, zu aufwändig, zu deutsch" (Reppner 2010) diskutiert und überfordern viele Unternehmen (Vgl. Lange 2015). Hierdurch erfreut sich die ISO 27001 zunehmender Beliebtheit, was sich z.B. in einer Verdopplung der ISO 27001-zertifizierten Unternehmen in den Jahren 2009 bis 2013 zeigt (Vgl. International Organization for Standardization 2014). Die Norm gesteht den Unternehmen deutlich mehr Freiheitsgrade zu und erlaubt eine individuell, auf die jeweilige Unternehmenssituation zugeschnittene Ausgestaltung des ISMS. Allerdings mangelt es der Norm an praktischen Umsetzungshinweisen, sodass auch die ISO 27001-Vertreter gerne auf die GS-Standards zurückgreifen, um konkrete Arbeitsschritte und Methoden für die Ausgestaltung eines ISO 27001-konformen Vorgehens im Sinne eines kontinuierlichen Sicherheitsprozesses zu übernehmen (Vgl. Kersten et al. 2013).

Die so entstehenden "ISMS-Mischformen" weisen jedoch häufig methodische Brüche und Inkonsistenzen auf, da sich die Arbeitsschritte und Methoden des BSI-Sicherheitsprozess nicht ohne Weiteres in ein ISO 27001-konformes Vorgehen überführen lassen. Während einige Arbeitsschritte innerhalb der GS-Standards keine eindeutige Entsprechung in der Norm finden oder sich nur durch starke Interpretation zuordnen lassen, können andere gleich mehreren unterschiedlichen Anforderungen der ISO 27001 zugeordnet werden. Dieses Bild verdeutlicht auch die eigens durch den BSI bereitgestellte Mapping-Tabelle, die eigentlich eine mühelose Übertragbarkeit der GS-Standards im Allgemeinen und des BSI-Sicherheitsprozesses im Speziellen auf die ISO 27001 suggerieren soll (Vgl. BSI 2011). Bei genauerer Betrachtung der Mapping-Tabelle wird jedoch deutlich, dass diese sehr allgemein gehalten wurde und teilweise mehrere Kapitel unterschiedlicher GS-Standards sowie Elemente aus Maßnahmenkatalogen und Bausteinen einzelnen Anforderungen der Norm zuordnet werden mussten. Auch die Reihenfolge der Arbeitsschritte im BSI-Sicherheitsprozess und die damit implizit verbundene Priorisierung einzelner Aktivitäten sorgt für Verwirrung: so wird beispielsweise die Risikoanalyse gemäß GS-Standards erst dann initiiert, wenn ein sehr hoher Schutzbedarf festgestellt wurde. Im Gegensatz dazu ist das kontinuierliche Durchlaufen eines Risikobeurteilungsprozess eine der zentralen Anforderungen der ISO 27001 (Vgl. ISO/IEC 27001 2015).

Die diskutierten Aspekte stellen nur einige der Ungereimtheiten dar, die zwischen GS-Standards und ISO 27001 bestehen. Sie machen jedoch deutlich, dass eine Übertragung von Arbeitsschritten und Methoden nicht ohne Weiteres möglich und die Entwicklung eines, eigens für die ISO 27001 konzipierten Sicherheitsprozess erforderlich ist. Im Rahmen dieser Forschungsarbeit wird die Forschungslücke adressiert, indem sowohl induktiv (auf Basis von Beobachtungen in der ISMS-Praxis) als auch deduktiv (in Form von theoretischen Überlegungen zu den Anforderungen der ISO 27001) (Vgl. Becker 2000), ein idealtypischer, normkonformer Sicherheitsprozess als Referenzprozess entwickelt und in der Praxis evaluiert wird.

### 3 Der ISO27-Sicherheitsprozess

Bereits in den früheren Versionen der ISO 27001 wurde das ISMS-Vorgehen prozesshaft in einem typischen Plan-Do-Check-Act Zyklus dargestellt. Obwohl die aktuelle Norm diese Gliederung nicht mehr explizit enthält, lässt sich die Struktur implizit aus den Normanforderungen ableiten und soll als Basisstruktur für den neuen ISO27-Sicherheitsprozess übernommen werden.

Da sich nicht alle Anforderungen der Norm in prozessuale Arbeitsschritte übertragen lassen, versteht sich nachfolgende Prozessbeschreibung als Leitfaden, welcher nur unter Beachtung aller Normanforderungen auch tatsächlich zu einem ISO 27001-konformen Vorgehen führen kann. Die einzelnen Phasen und Arbeitsschritte des ISO27-Sicherheitsprozess werden in nachstehender Tabelle dargestellt: In der Spalte "Aktion" findet sich eine Kurzbeschreibung des durchzuführenden Arbeitsschrittes, die Spalte "Erläuterung" enthält Ausführungen zur Zielsetzung des Arbeitsschrittes und die Spalte "ISO" verweist auf die entsprechende(n) Norm-Anforderung(en), die mit dem jeweiligen Arbeitsschritt adressiert werden.

Aktion	Erläuterung	ISO
<b>Plan-Phase:</b> Im Rahmen der Planungsphase werden die Grundsteine für jeden neu gestarteten Zyklus des ISO27-Sicherheitsprozess gelegt. Alle ISMS-relevanten Ziele, Vorgaben und operativ anzuwendenden Methoden sind zu definieren, bzw. zu aktualisieren. Es können strategische, taktische und operative Planungsschritte unterschieden werden, die jeweils sequentiell zu durchlaufen sind. Innerhalb einer Planungsphase kann die Bearbeitung der Arbeitsschritte auch parallel erfolgen.		
Strategische Planung		
Schritt 1: Definieren/ Aktualisieren Sie die Informations- sicherheitsziele und die ISMS-Politik	In der Phase der strategischen Planung des ISO27-Sicherheitsprozess sollte sich die Organisation einen Überblick über die aktuelle Situation zur Informationssicherheit verschaffen, um die Informationssicherheitsziele und die ISMS-Politik definieren/ aktualisieren zu können. Wird der ISO27-Sicherheitsprozess bereits wiederholt durchlaufen, können dabei auch Informationen aus vorherigen Zyklen Berücksichtigung finden.	4.4 5.1 5.2 6.2 7.5
Taktische Planung		
Schritt 2.1: Definieren/ Aktualisieren Sie den ISMS-Anwendungs- bereich (Scoping)	Der ISMS-Anwendungsbereich definiert, welche Geschäftsprozesse (GP) und welche, an deren Bearbeitung beteiligte Organisationseinheiten (OE'n) operativ in das ISMS eingebunden werden sollen. Hierzu ist zunächst der Kontext der Organisation, z.B. Aufbau- und Ablauforganisation zu erfassen. Die Auswahl der relevanten Bereiche ist unternehmensspezifisch zu treffen und in einem Scope-Dokument zu dokumentieren.	4.1 4.2 4.3 4.4 5.1 7.5
Schritt 2.2: Definieren/ Aktualisieren Sie die ISMS-Organisation und überprüfen Sie die Ressourcen- planung	Die ISMS-Organisation sollte parallel zur, oder nach der Festlegung des Anwendungsbereiches definiert werden, da dieser bestimmt, welche ISMS-Rollen für den ISO27-Sicherheitsprozess erforderlich sind und welche Rolleninhaber infrage kommen. Während der Planungsperiode sollte außerdem definiert/ überprüft werden, ob die ISMS-Organisation angemessen mit Ressourcen ausgestattet ist, die Rollenbeschreibungen aktuell und die zugewiesenen Rolleninhaber korrekt benannt sind. Weiterhin sind Qualifikationspläne für die Rollen zu erarbeiten/ aktualisieren und ggf. erforderliche Qualifikationsmaßnahmen einzuleiten.	4.4 5.1 5.3 7.1 7.2 3.4. 7.2
Operative Planung		
Schritt 3.1: Definieren/ Aktualisieren Sie die Vorgaben zur Dokumentenlenkung	Alle ISMS-relevanten Dokumente sollten gemäß Norm angemessen gekennzeichnet sein, hinsichtlich möglicher Änderungen überwacht und vor Veröffentlichung geprüft und genehmigt werden. Die hierfür erforderlichen Prozesse und Vorgaben sollten im Rahmen der operativen Planungsphase definiert/ aktualisiert werden.	4.4 7.5
Schritt 3.2: Planen/ Aktualisieren Sie den	Der Risikomanagementprozess ist ein zentraler Teilprozess des ISO27-Sicherheitsprozess, dessen Ablauf und zugrundeliegende Methoden	4.4 6.1

Aktion	Erläuterung	ISO
Risikomanagementprozess	detailliert geplant werden sollten. Beispielsweise sind Methoden für die Risikoidentifizierung und -behandlung festzulegen, Risikobehandlungsstrategien zu bestimmen und die am Risikomanagement beteiligten Mitarbeiter zu benennen. Außerdem sollten alle relevanten Vorgaben und Prozessbeschreibungen in einem zentralen Dokument zum Risikomanagement dokumentiert werden.	
Schritt 3.3: Erstellen/ Aktualisieren Sie Schulungs- und Kommunikationskonzepte	Ebenfalls im Rahmen der operativen Planungsphase sollten die Schulungskonzepte für die Belegschaft sowie die Kommunikations- und Eskalationskonzepte erarbeitet und in der ISMS-Dokumentation festgehalten werden.	4.4 5.1 7.3 7.4
Schritt 3.4: Definieren/ Aktualisieren Sie Überwachungsprozesse und planen Sie interne Audits	In Vorbereitung auf die Check-Phase sollten bereits zum Planungszeitpunkt die Überwachungs- und Messverfahren sowie Vorgaben und Zeitpläne für interne Audits definiert werden, um später die Wirksamkeit des ISMS überprüfen zu können. Auch das Berichtswesen sollte bereits jetzt konzipiert werden.	4.4 5.1 9.1 9.2 9.3
<b>Do-Phase:</b> Im Rahmen der Do-Phase sind insbesondere die im Rahmen der operativen Planung festgelegten Teilprozesse zu initiieren und geplante Maßnahmen zu ergreifen. Die nachfolgend beschriebenen Arbeitsschritte können, je nach Ressourcensituation auch parallel abgearbeitet werden.		
Schritt 4.1: Initiieren/ Starten Sie den Risikomanagementprozess.	Gemäß ISO 27001 sind für alle, im Anwendungsbereich definierten GP Informationssicherheitsrisiken zu erfassen und zu steuern. Die Methodik dafür wurde im Rahmen der operativen Planung festgelegt (Vgl. Schritt 3.2). Sollte der ISO27-Sicherheitsprozess bereits wiederholt durchlaufen werden, können die Ergebnisse aus dem vorherigen ISMS-Zyklus als Informationsquelle hinzugezogen werden.	4.4 8.1 8.2
Schritt 4.2: Setzen Sie definierte IS-Maßnahmen um.	Im Ergebnis des Teilprozesses zum Risikomanagement werden meist eine Vielzahl von Schutzmaßnahmen definiert, die es umzusetzen gilt. Die Maßnahmen dienen i.d.R. dazu, bestehende technische oder organisatorische Schwachstellen innerhalb der Organisation zu beseitigen. Da die Fertigstellung der Maßnahmen meist einige Zeit in Anspruch nimmt, sollte ein zentrales Maßnahmentracking aufgebaut werden.	4.4 8.1 8.3
Schritt 4.3: Initiieren Sie Schulungen und Awareness-Maßnahmen	Da nicht alle Mitarbeiter der Organisation direkt im Rahmen des ISO27-Sicherheitsprozess mitwirken aber dennoch einen Beitrag zur Informationssicherheit leisten können und müssen, sind regelmäßig Schulungen und Awareness-Maßnahmen zu initiieren, um die Mitarbeiter für das Thema zu sensibilisieren (Vgl. Schritt 3.3)	4.4 5.1 8.1 7.3
<b>Check-Phase:</b> Im Rahmen der Check-Phase werden Wirksamkeit des ISMS und Regelkonformität zu den ISMS-Vorgaben überwacht und bewertet. Die einzelnen Arbeitsschritte in dieser Phase müssen nicht sequentiell abgearbeitet sondern können parallel durchlaufen werden.		
Schritt 5.1: Überprüfen und Bewerten Sie die Einhaltung/ Umsetzung der ISMS-relevanten Vorgaben.	Damit das ISMS seine Wirkung entfalten und zur Steigerung der operativen Informationssicherheit beitragen kann, ist regelmäßig zu überprüfen, ob die definierten Vorgaben und Teilprozesse des ISO27-Sicherheitsprozess eingehalten und gelebt werden. Die jeweiligen Messpunkte, -methoden und -zyklen wurden im Rahmen der operativen Planung festgelegt (Vgl. Schritt 3.4) und die Ergebnisse sind zu dokumentieren.	4.4 5.1 8.1 9.1
Schritt 5.2: Führen Sie interne Audits durch.	Im Gegensatz zu den im vorherigen Punkt beschriebenen Messungen, welche i.d.R. durch Rolleninhaber des ISO27-Sicherheitsprozess durchgeführt werden, ist bei den internen Audits darauf zu achten, dass die Auditoren objektiv und unparteilich vorgehen. Es empfiehlt sich daher Mitarbeiter auszuwählen, die nicht direkt am ISO27-Sicherheitsprozess mitwirken oder von diesem betroffen sind.	4.4 5.1 8.1 9.2

Aktion	Erläuterung	ISO
Schritt 5.3: Erstellen Sie das Berichtswesen.	Die Ergebnisse der internen Wirksamkeits- und Konformitätsüberprüfungen sind in einem Berichtswesen (Vgl. Schritt 3.4) aufzubereiten und als Grundlage für die Managementbewertung zu nutzen.	4.4 8.1 9.3
Act-Phase: Im Rahmen der Act-Phase werden festgestellte Mängel beseitigt und Verbesserungsmaßnahmen für den nachfolgenden Zyklus des ISO27-Sicherheitsprozess angestoßen.		
Schritt 6.1: Leiten Sie Korrekturmaßnahmen bei Nicht-Konformität ein.	Die im Rahmen der Check-Phase festgestellten Mängel und Abweichungen sind durch das Ergreifen von Korrekturmaßnahmen zu beseitigen. Die Maßnahmen sind hinsichtlich ihrer Wirksamkeit zu überprüfen und zu dokumentieren.	4.4 5.1 8.1 10.1
Schritt 6.2: Initiieren Sie Verbesserungsmaßnahmen.	Verbesserungsvorschläge, die zu einem effizienteren oder effektiveren ISO27-Sicherheitsprozess und damit zu einer steigenden Informationssicherheit beitragen, sind zu erfassen und in nachfolgenden Prozesszyklen zu berücksichtigen.	4.4 5.1 8.1 10.2

## 4 Zusammenfassung und Forschungsausblick

Im vorliegenden Beitrag wurde der ISO27-Sicherheitsprozess als ein neuer Referenzprozess zur Umsetzung der Anforderungen aus der ISO 27001 vorgestellt. Der Prozess folgt dem PDCA-Modell, definiert phasenspezifische und konkrete ISMS-relevante Arbeitsschritte und bringt diese in eine sequentielle Reihenfolge. Zielsetzung ist es, die Normkonformität des Referenzprozesses sicherzustellen, sodass Unternehmen, die dem ISO27-Sicherheitsprozess folgen, "by-design" ein zertifizierungsfähiges ISMS aufbauen und betreiben.

Dem gestaltungsorientierten Forschungsparadigma der Wirtschaftsinformatik folgend (Vgl. Österle et al. 2010), soll der ISO27-Sicherheitsprozess in weiteren Forschungsiterationen gemeinsam mit der wissenschaftlichen Community und Praxispartnern diskutiert, verfeinert und im Realsystem evaluiert werden. Anschließend lassen sich die bislang sehr generisch beschriebenen Aktivitäten des ISO27-Sicherheitsprozess mit konkreten Umsetzungshinweisen aus den GS-Standards untersetzen, um Synergien zwischen beiden ISMS-Standards zu heben, ohne dabei die Normkonformität des Referenzprozess zu gefährden. Darüber hinaus soll der ISO27-Sicherheitsprozess als Basis für ein speziell für den ISMS-Betrieb konzipiertes Workflow-Management-System namens CISO27 dienen, mithilfe dessen die Teilprozesse und Arbeitsschritte des Referenzprozesses innerhalb von Unternehmen IT-seitig unterstützt und (teil-) automatisiert werden können. Hierdurch soll die Einbindung dezentraler Sicherheits- und Risikoverantwortlicher unterstützt werden.

## 5 Literatur

- Becker J (2000) Informationsmodelle für das Electronic Business. In: Weiber R (Hrsg) Handbuch Electronic Business: Informationstechnologien — Electronic Commerce — Geschäftsprozesse. Gabler, Wiesbaden
- BSI (2015) IT- Grundschatz. <https://www.bsi.bund.de>. Abgerufen am 26.09.2015
- BSI (2011) Zuordnungstabelle ISO 27001 sowie ISO 27002 und IT-Grundschatz (aktualisiert). [https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Grundschatz/Hilfsmittel/Doku/Vergleich\\_ISO27001\\_GS.pdf](https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Grundschatz/Hilfsmittel/Doku/Vergleich_ISO27001_GS.pdf). Abgerufen am 20.09.2015



- International Organization for Standardization (2014) The ISO Survey of Management System Standard Certifications. <http://www.iso.org/iso/iso-survey>. Abgerufen am 15.08.2015
- ISO/IEC 27001 (2015) Informationstechnik – IT-Sicherheitsverfahren – Informationssicherheits- Managementsysteme – Anforderungen
- Kersten H, Reuter J, Schröder KW (2013) IT-Sicherheitsmanagement nach ISO 27001 und Grundschatz: Der Weg zur Zertifizierung. 4.Auflage. Springer Vieweg, Wiesbaden
- Klipper S (2015) Information Security Risk Management: Risikomanagement mit ISO/IEC 27001, 27005 und 31010. 2.Auflage. Springer Vieweg, Wiesbaden
- Lange J (2015) Die (wirkliche) Modernisierung des IT-Grundschatz. <http://it-sibe.de/2015/01/die-wirkliche-modernisierung-des-grundschatz>. Abgerufen am 27.09.2015
- Österle H (Hrsg), Winter R (Hrsg), Brenner W (Hrsg) (2010) Gestaltungsorientierte Wirtschaftsinformatik: Ein Plädoyer für Rigor und Relevanz. Infowerk, Nürnberg
- Reppner M (2010) BSI-Grundschatz: zu komplex, zu aufwändig, zu deutsch. <http://www.zdnet.de/41525300/bsi-grundschatz-zu-komplex-zu-aufwaendig-zu-deutsch>. Abgerufen am 03.09.2015



# Effects of Performance Transparency in E-Learning Applications: A Field Study in Primary Schools

Martin Huschens<sup>1</sup>, Henning Müller<sup>2</sup>, Franz Rothlauf<sup>1</sup>, and Daniel Schunk<sup>2</sup>

<sup>1</sup> Universität Mainz, Lehrstuhl für BWL und Wirtschaftsinformatik, huschens@uni-mainz.de, rothlauf@uni-mainz.de

<sup>2</sup> Universität Mainz, Lehrstuhl für Public Economics, henning.mueller@uni-mainz.de, daniel.schunk@uni-mainz.de

## Abstract

The diffusion of (gamified) e-learning applications in primary schools is advancing fast. One key feature of such systems is that they increase performance transparency which facilitates social comparison processes between pupils. Yet the impact of these processes are not well understood. Thus the aim of our research project is to gain insights into the effects of such e-learning applications on motivation and learning outcomes. Furthermore, we are interested in the impact on children's risk preferences, social behaviour and affective outcomes (perceived pressure to perform). We use a randomized field study in elementary schools in the city of Mainz to compare e-learning systems that differ with regard to their performance transparency. Results will help to develop guidelines for the development and usage of educational software in the classroom.

## 1 Introduction

Teaching the handling of technological devices is a key function of school education in Germany. In the wake of this directive, the diffusion of e-learning applications already in primary schools has rapidly been increasing. Pilot projects integrate digital learning devices into the school routines either based on technological devices like tablets and e-book readers or using software applications like Moodle (e-learning platform) or Antolin (platform to increase reading skills).

A partly intended side effect of such systems is their increased performance transparency. This can be caused by two processes: either the classmate's performance is revealed implicitly by the possibility to observe actions and learning processes of peers on an (online) learning platform. Or a peer's performance may also be presented explicitly in the form of public leaderboards. This is of special importance in the realm of gamification that tries to implement game design elements like levels, badges or rankings into non-gaming contexts (Blohm and Leimeister 2013). As a consequence, especially gamified e-learning systems not only expose pupils' learning results but also their learning process which leads to increased social comparison (Festinger 1954). Nevertheless, the impact of such systems that induce increased social comparison processes are not well understood.

To gain insights into processes of social comparison in the field, we conduct a randomized field study at primary schools in the city of Mainz. We apply a treatment and control group design in which we vary performance transparency by implementing different social comparison sources into an e-learning setting (treatment group: class wide leaderboard, control group: only individual score). The e-learning application will be used within schools during math lessons. Our research will be guided by the following research question: What are the effects of performance transparency in e-learning applications on pupils' motivation, learning outcomes, social behaviour, risk preferences and perceived pressure?

The paper is structured as follows: section two presents the theoretical background on e-learning, performance transparency and resulting social comparison processes by providing a literature review. In section three we develop our research model and hypotheses. Finally we describe the design of our field study. Our experimental design will be based on a power analysis. Firstly, we present the results of the analysis to derive a grounded sampling and randomization strategy and to further estimate the expected treatment effects. Lastly, we present the tools for data collection.

## 2 Theoretical Background & Literature Review

### 2.1 E-Learning, Gamification and Performance Transparency

E-learning describes the fact of integrating information systems into the learning process. Clark and Mayer (2011) define e-learning as “instruction delivered on a digital device that is intended to support learning” (Clark and Mayer 2011). In addition, we understand e-learning applications not only as technological artefacts that transfer offline content into a digital context but also as a pedagogical method based on didactic principles using information systems (Bronfman 2000).

In turn, such didactic principles are based on learning theories that describe how individuals acquire knowledge. Some of the most influential learning theories state that learning is based on social interaction. Humans observe the (learning) behaviour and outcomes of peers and imitate this behaviour. Furthermore, humans compare their results with peers and evaluate them in relation to others (social learning theory, Bandura 1971; socio cultural learning theory, Vygotsky 1978). This notion of creating a social surrounding to improve learning outcomes is also implemented in e-learning applications that use *gamification* elements to provide social comparison information.

Kapp and colleagues (2014) define gamification in the learning context as “using game-based mechanics, aesthetics, and game-thinking to engage people, motivate action, promote learning, and solve problems” (Kapp et al. 2014, p. 125). They distinguish between *structural gamification* and *content gamification*. Whereas in the case of *content gamification*, contents are altered to create a game-like experience, e. g. by adding story elements or transforming content to simulation games, *structural gamification* adds game elements like points, leaderboards or badges with no changes to the content. In order to create motivational effects, *structural gamification* relies, among others, on one important fundamental: *transparency* (Kapp et al. 2014). Gamified e-learning applications provide pupils with information about peer's performance and learning processes in order to facilitate social comparison processes and to foster motivation (other outcomes and their definition can be found in table 1). Further, such performance transparency might be the basis of gaining a social status, e. g. by being the leader on a leaderboard (Kapp et al. 2014).

## 2.2 Social Comparison Effects

Social comparison has been a vital research area for decades. Already in 1954, Festinger presented the foundation of this theory. His insights are based on two fundamental assumptions: First, humans have a basic desire to compare themselves with others and, second, they have a basic drive upwards towards higher performance. As a consequence social comparison leads to competitive behaviour which finally may have a positive impact on performance and motivation (Festinger 1954). A comprehensive literature review of social comparison in theory and research can be found in Suls and Wheeler (2000).

Dijkstra and colleagues (2008) present a review on social comparison processes in educational settings. Results are dichotomous: Performance transparency as such is not always beneficial for every individual. Empirical studies point to detrimental social comparison effects on performance and motivation, especially for those subjects who constantly experience inferiority. Low performers often perceive additional pressure from a situation of being forced to perform in order show superior results. This process can even further inhibit their performance (*choking under pressure*, Baumeister and Showers 1986). That is why Dijkstra and colleagues term social comparison in the classroom a “two-edged sword” (Dijkstra et al., 2008, p. 865). However, those studies did not take into account gamified e-learning applications. Plass and colleagues (2013) provide a first approach and found positive effects on performance, motivation and interest. Yet, they neither look at individual consequences nor did they study effects on social behaviour and risk preferences.

However, it is likely that social comparison in the classroom may also have effects on pupil’s social behaviour and willingness to cooperate. Especially gamified e-learning applications that incorporate a leaderboard entail a competitive situation for every pupil. That is because pupils’ goal attainment is negatively correlated: a higher position for one subject is always linked to a lower rank of another subject which equals a loss concerning the social status (Garcia et al. 2013).

Further, the provision of social comparison information (e. g., by leaderboards) may also lead to a shift in pupils’ risk preferences. This is due to the fact that the gain in social situations is twofold: on the one hand pupils receive points for their learning outcomes and on the other hand they gain a certain social status by achieving a better ranking. We assume that pupils will favour more difficult tasks in order to gain more points and, consequently, exposing themselves to greater possibilities of negative outcomes (wrong answers). Research on this outcome variable in the educational domain is very sparse. However, research in the economic and management literature proves the fundamental link between performance transparency and risk preferences: Social comparison information fosters risk affinity, e. g. concerning investment decisions (Dijk et al. 2014).

Outcomes	Definition
Motivation	Internal processes that give behaviour its energy and direction (Ryan and Deci 2000)
Learning outcomes	Achieved result in a math test (difference-in-differences measurement based on a baseline measurement)
Perceived pressure	Perceived presence of situational incentives for optimal, maximal or superior performance often in the form of competition (Baumeister and Showers 1986)
Risk preferences	Willingness to expose oneself to more difficult tasks or to the possibility of loss/negative outcomes (Schonberg et al. 2011)
Social behaviour/ cooperation	Behaviour that is based on cooperation and motivated by altruistic motives rather than being self-interested (measured in a public goods game, Andreoni 1988)

**Table 1: Construct definitions**

### 3 Research Model & Hypotheses

To answer our research question we conduct a randomized field study within primary schools in the city of Mainz. In each school we will introduce an e-learning application for third grade classes. Approximately 400 pupils in 20 school classes will use the system for some weeks (treatment phase). We implement a blended learning approach in which the application will be embedded into regular math lessons as additional digital revision of the content so that pupils will use the system various days in the week. During the treatment phase, pupil's performance will be tracked which will be the basis for leaderboards and consequently performance transparency.

We implement a treatment and control group design. The two groups vary with regard to the implemented performance transparency (from high to low): The treatment group receives a class wide leaderboard with a cardinal scale so that pupils can both see their position relative to everyone else and their absolute score (full performance distribution). The control group will have a low level of performance transparency as no leaderboard will be delivered to the pupils. Nevertheless, as in the treatment group, their own score will be displayed which allows for comparison outside the e-learning application. So even without performance transparency in the e-learning application social comparison processes can never be ruled out entirely but at least be minimized. However, to account for the effect of mere feedback on performance, the control group will also receive information on individual performance (without relative ranking information).

Our research model can be seen in figure 1. The model posits that the level of performance transparency (high and low) will directly affect the outcome variables. In general, our hypotheses are structured in two areas: The first area makes assumptions about the effect of performance transparency in e-learning applications on the class level. Within area two we also consider the individual level. In particular, we hypothesize a moderating effect of the individual ranking position on the relationship between the level of performance transparency and motivation, perceived pressure and risk preferences in the direction stated in figure 1.

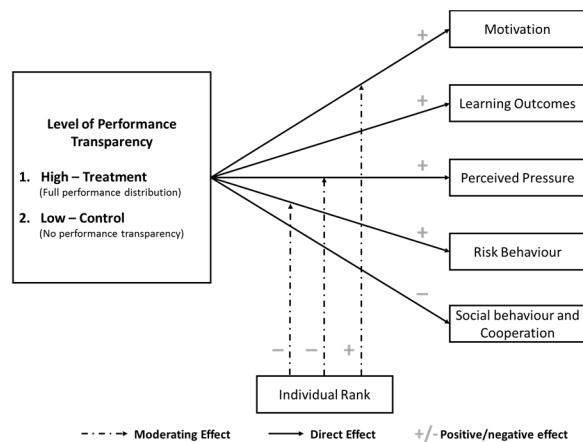


Figure 1: Research model

## 4 Experimental Design

### 4.1 Power Analysis

Identifying the appropriate experimental design for a study is critical because of limited resources. Hence, our experimental design is based on a power calculation following the *effect size approach* (Gerber and Green 2012): the aim is to detect the minimum detectable effect size (MDES) at a given sample size to achieve a desired level of statistical power. Firstly, this allows us to define the scientific risk related to our field study and, secondly, to design our study to maximize statistical power.

Based on a recent working paper by Goulas and Megalokonomou 2015 who examined the effects of social comparison information on student's performance and subsequent university enrolment in Greece, we can estimate the effect size of our treatment to be approximately  $d = 0.25$  to  $0.35$  (as our treatment is much more intensive, we expect our effect sizes to be rather larger than in Goulas and Megalokonomou (2015)). A power analysis with the software "Optimal Design" (Raudenbush et al. 2011) based on these effect sizes and using a blocked design (blocks ( $K = 10$ ) are schools, clusters ( $J = 2$ ) are classes) results in power estimations displayed in Figure 2. Effect sizes of  $d = 0.25 / 0.30 / 0.35$  on our outcome variables will be detected on a 5% significance level with a power of 52,2% / 67,5% / 80,2%, respectively.

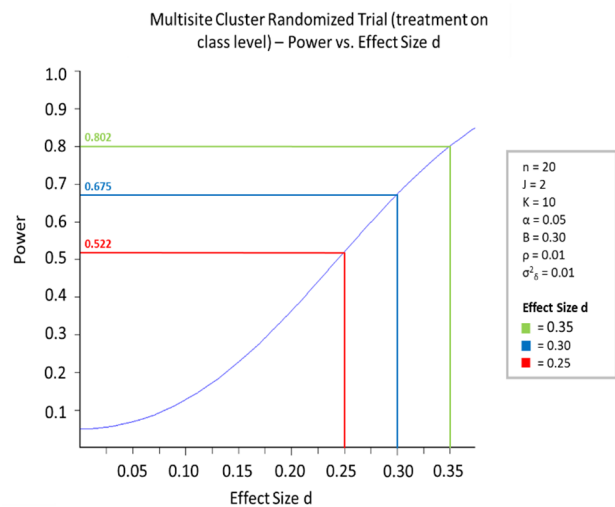


Figure 2: Results of power analysis

Nevertheless, as we do not have the possibility to increase our sample size significantly because of a limited number of cooperating schools we will have to ensure that the effect size is sufficiently large and that pre-treatment noise influencing the outcomes is reduced. Thus, we administer two methods to further enlarge the power: a randomization strategy that uses optimized blocking and the difference-in-differences measurement design (see next sections).

#### 4.2 Randomization Strategy

In order to allocate classes to the treatment and control group we use a block random assignment on cluster level. Following this design randomization takes place on school class level which means that we assign subjects (pupils) not as an individual unit but rather as cluster (school classes) to the experimental groups. Although this reduces the independent observations significantly this solution is the only reasonable study design for two reasons: firstly, only the randomization on class level allows us to separate the treatments. Intra-class randomization would lead to a dilution of the treatments and spill-over effects based on communication between pupils. Secondly, this allows us to conduct the treatments for a longer period of time (about 4 weeks), which will in consequence lead to larger effects on the outcome variables.

Further, we implement a block design in which we divide our clusters into 10 subgroups called **blocks or strata, such that the** variability within blocks is less than the variability between blocks. Classes within each stratum are allocated randomly to treatment and control group. This assures that the probability of assignment to the treatment is constant in each block. By forming these homogenous blocks (based on the socio-economic status) we substantially reduce pre-treatment noise. Consequently, by controlling for relevant covariates already in the randomization process we improve precision and power of our experimental study significantly.

#### 4.3 Data Collection Methods

We apply several data collection methods to measure the treatment effects on the outcome variables. In order to meet the requirements of our subjects, we implement child-oriented data collection

methods mainly based on games. Methodically we use a *difference-in-differences measurement approach* by conducting a baseline measurement and a post-treatment measurement. On the one hand, this increases precision of our estimates by reducing the variance in the outcome. On the other hand, we further reduce the selection bias as differences between the two groups' tests scores are more likely to *purely* reflect the effect of the manipulation and are not a consequence of absolute class differences. Additionally, as our data will be organized in more than one level (nested data) we use a multilevel analysis. This allows us to separately estimate the influence of performance transparency on the outcomes for pupils within a class and between classes.

Motivation is a multi-dimensional construct which is comprised of several sub constructs (Ryan and Deci 2000). Thus, on the one hand, we apply behavioural measures of motivation, e. g. by observing pupil's speed in the math game. Additionally, we give pupils the opportunity to engage themselves in a leisure activity (e. g. watching a cartoon) after having completed their obligatory math exercises. This allows us to identify pupils who deliberately decide for the math game and show higher motivation than others. On the other hand, we will also need to use a questionnaire to measure cognitive and affective sub facets of motivation. The questionnaire will capture how the treatment affects intrinsic (learning fun, interest, enjoyment etc.) and extrinsic motivation (based on the Intrinsic Motivation Inventory, Ryan and Deci 2000). Further, the questionnaire will comprise questions concerning the perceived pressure by the pupil's.

Learning outcomes will be measured within the math game and are equivalent to the pupil's score. Furthermore, we will conduct a pre-treatment and post-treatment math test that allows us to measure the trend in the learning outcomes, apart from task learning effects.

Pupil's risk preferences will be measured with two risk tasks. Both games measure behavioural risk attitudes of subjects. Whereas in the Balloon Analogue Risk Task (BART, Lejuez et al. 2002) subjects earn points by inflating a balloon without knowing when it will explode (which leads to the loss of all gains), the Devil's Task (Slovic 1966) entails the task to open treasure chests without knowing in which chest the devil is hidden that will take all the winnings. The measures of risk preferences are the average number of pumps or respectively opened chests.

Finally, pupils' social behaviour and willingness to cooperate will be measured in a version of the public goods game (Andreoni 1988) that is adapted to the cognitive skills of children. The game involves the task to decide how much of an initial endowment one is willing to invest in a public good that yields a specific return for every participant. Even those subjects who did not invest in the public good will participate in the returns. The game allows to identify free riding behaviour of pupils which we can be interpreted as a proxy for their willingness to cooperate.

## 5 Outlook

After the conceptualization of the study and approved funding by the BMBF, our project officially kicked-off in August 2015. We already received approval from an ethical commission and are now applying for a formal approval from the responsible school authorities. Currently, we are also developing and technically implementing the learning software. In the next phase, we will recruit participating schools and prepare all the field work. The intervention period is planned for summer/autumn 2016.



## 6 Bibliography

- Andreoni J (1988) Why Free Ride? Strategies and Learning in Public Goods Experiments. *J Public Econ* 37(4):291–304
- Bandura A (1971) *Social Learning Theory*. General Learning Press, New York
- Baumeister RF, Showers CJ (1986) A Review of Paradoxical Performance Effects: Choking Under Pressure in Sports and Mental Tests. *Eur J Soc Psychol* 16(4):361–383
- Blohm I, Leimeister JM (2013) Gamification. *Wirtschaftsinformatik* 55(4):275–278
- Bronfman SV (2000) Linking Information Technology and Pedagogical Innovation to Enhance Management Education. In: *Proc. 8th ECIS 2000*. Vienna, Austria, pp 1405–1412
- Clark RC, Mayer RE (2011) *E-Learning and the Science of Instruction*, 3 rd. Wiley, San Francisco
- Dijk O, Holmen M, Kirchler M (2014) Rank Matters - The Impact of Social Competition on Portfolio Choice. *Eur Econ Rev* 66(2014):97–110
- Dijkstra P, Kuyper H, Werf G Van Der, et al (2008) Social Comparison in the Classroom : A Review. *Rev Educ Res* 78(4):828–879
- Festinger L (1954) A Theory of Social Comparison Processes. *Hum Relations* 7(2):117–140
- Garcia SM, Tor A, Schiff TM (2013) The Psychology of Competition: A Social Comparison Perspective. *Perspect Psychol Sci* 8(6):634–650
- Gerber AS, Green DP (2012) *Field Experiments: Design, Analysis and Interpretation*. Norton, NY
- Goulas S, Megalokonomou R (2015) Knowing who you are: The Effect of Feedback Information on Short and Long Term Outcomes. *Warwick Economics Research Paper Series*
- Kapp KM, Blair L, Mesch R (2014) *Gamification of Learning and Instruction Fieldbook: Ideas into Practice*. Wiley, San Francisco
- Lejuez CW, Read JP, Kahler CW, et al (2002) Evaluation of a Behavioral Measure of Risk Taking: The Balloon Analogue Risk Task (BART). *J Exp Psychol Appl* 8(2):75–84
- Plass JL, O’Keefe P a., Homer BD, et al (2013) The Impact of Individual, Competitive, and Collaborative Mathematics Game Play on Learning, Performance, and Motivation. *J Educ Psychol* 105(4):1050–1066
- Raudenbush SW, Spybrook J, Martinez A, et al (2011) *Optimal Design Software for Multi-level and Longitudinal Research (V. 3.01) Software*. Available from [www.wtgrantfoundation.org](http://www.wtgrantfoundation.org)
- Ryan RM, Deci EL (2000) Self-determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *Am Psychol* 55(1):68–78
- Schonberg T, Fox CR, Poldrack R a. (2011) Mind the gap: Bridging economic and naturalistic risk-taking with cognitive neuroscience. *Trends Cogn Sci* 15(1):11–19
- Slovic P (1966) Risk-Taking in Children: Age and Sex Differences. *Child Dev* 37(1):169–176
- Suls J, Wheeler L (2000) *Handbook of Social Comparison: Theory and Research*. Kluwer Academic Publishers, Dordrecht
- Vygotsky LS (1978) *Mind in society: The development of higher psychological processes*. Harvard University Press, Cambridge, MA



# Vom Standard zur Anwendung: Ein Blick in Syntax, Semantik und Pragmatik der Adaption von BPMN

Florian Imgrund<sup>1</sup> und Christian Janiesch<sup>1</sup>

<sup>1</sup> Julius-Maximilians-Universität Würzburg, Juniorprofessur für Information Management, {florian.imgrund, christian.janiesch}@uni-wuerzburg.de

## Abstract

Die Business Process Model and Notation (BPMN) hat sich seit ihrer Einführung vom de-facto Standard zu einem internationalen ISO-Standard der heutigen Prozessmodellierung entwickelt. Im Sprachgebrauch der BPMN wird zwischen Syntax, Semantik und Pragmatik unterschieden. Die angestellte Literaturanalyse zeigt unter der Einführung eines Ebenen-Modells für die Prozessmodellierung, dass die in den offiziellen Spezifikationsdokumenten deutliche Vernachlässigung der pragmatischen Komponente der Sprache durch eine hohe Publikationsdichte entsprechender Fachliteratur in weiten Teilen behoben wurde. Die Literaturanalyse zeigt jedoch auch, dass pragmatische Hilfestellungen erst bei der konstruktiven und analytischen Evaluation von Prozessmodellen einsteigen. Der dem eigentlichen Modellierungsprozess vorgelagerte Aspekt der Entwicklung eines mentalen Modells und der damit verbundenen kognitiven Belastung des Modellierers noch *vor* der Modellierung wird aber aktuell noch unzureichend berücksichtigt.

## 1 Einleitung

Das Geschäftsprozessmanagement (engl.: Business Process Management, BPM) hat bereits eine Reihe verschiedener Modellierungstechniken hervorgebracht, durch die sich dessen Stakeholder eine Verbesserung des Verständnisses für Geschäftsprozesse, eine Erhöhung der Prozessqualität als auch eine vereinfachte Kommunikation sowohl im als auch zwischen Unternehmen versprechen (Indulska et al. 2009a; Overhage et al. 2012). Die graphische Modellierungssprache Business Process Model and Notation (BPMN) (Object Management Group 2013) hat sich seit der Einführung 2003 durch die Business Process Management Initiative (BPMI) vom de-facto Standard zum internationalen ISO-Standard der heutigen Prozessmodellierung durchgesetzt (Owen und Raj 2003; Recker 2010). Zum gegenwärtigen Zeitpunkt liegt die Notation in einer zweiten, überarbeiteten Version vor und wird von der Object Management Group (OMG) verwaltet. Sowohl in der Entwicklung durch die BPMI als auch nach der Übernahme des Standards durch die OMG wird die Notation von einem non-Profit Konsortium betreut, das aus einem breiten Spektrum verschiedener IT-Unternehmen besteht.

In der Literatur steht die wertschöpfende Wirkung der BPMN im Kontext des Managements von Geschäftsprozessen außer Frage (Muenstermann et al. 2010; Indulska et al. 2009a). Dennoch

weisen zahlreiche Publikationen im Bereich der evaluativen Bewertung des Einsatzes von BPMN auf Anwendungsebene auf eine suboptimale Adaption des Standards in der Praxis hin (Leopold et al. 2015; Haisjackl et al. 2015; Recker 2010). Ausschlaggebend sind hierfür ausbleibende Hilfestellungen in Form pragmatischer Handlungsanweisungen oder die fehlende Bereitstellung bewährter Vorgehensweisen innerhalb des Spezifikationsdokuments (zur Muehlen et al. 2010). Recker (2010) und Leopold et al. (2015) zeigen, dass sich diese Entwicklung unabhängig der Verfügbarkeit verschiedener Qualitätsframeworks wie beispielsweise den Guidelines of Modeling (Schuette und Rotthowe 1998), dem Semiotic Quality Framework (Krogstie et al. 2006) oder den Seven Process Modeling Guidelines (Mendling et al. 2010) fortzuführen scheint. So zählt die *Modellierung von Subprozessen mit inkonsistenten Verbindungen zum jeweiligen Hauptprozess* die Erstellung von *Modellen übermäßiger Diagrammgröße* oder die *fehlende Verlinkungen eingesetzter Datenobjekte* noch immer zu den am häufigsten beobachteten Fehlern in der Prozessmodellierung (Leopold et al. 2015). Die erfolgreiche Anwendung einer umfangreichen und weit entwickelten Notationssprache wie der BPMN setzt ein ausgebildetes Verständnis sowohl für die Modellierungsgrammatik (Syntax und Semantik) als auch dessen pragmatische Anwendung voraus (Muenstermann et al. 2010). Die anhaltende Fehlerhaftigkeit der modellierten Geschäftsprozesse impliziert, dass das in der Praxis angewendete Vorgehen zur Prozessmodellierung noch nicht seinen letzten Reifegrad erreicht hat. In diesem Zusammenhang sollen folgende Fragestellungen untersucht werden:

1. *Bedient die OMG mit der BPMN die Anforderungen der Praxisanwender hinsichtlich Syntax und Semantik sowie der Umsetzung auf pragmatischer Ebene?*
2. *Welche Rolle spielen weiterführende Dokumente (Handlungsanleitungen, Governance-Dokumenten, Konventionen etc.) in der Adaption und Nutzung der BPMN? Decken diese den gesamten Prozess der Modellierung zufriedenstellend ab?*

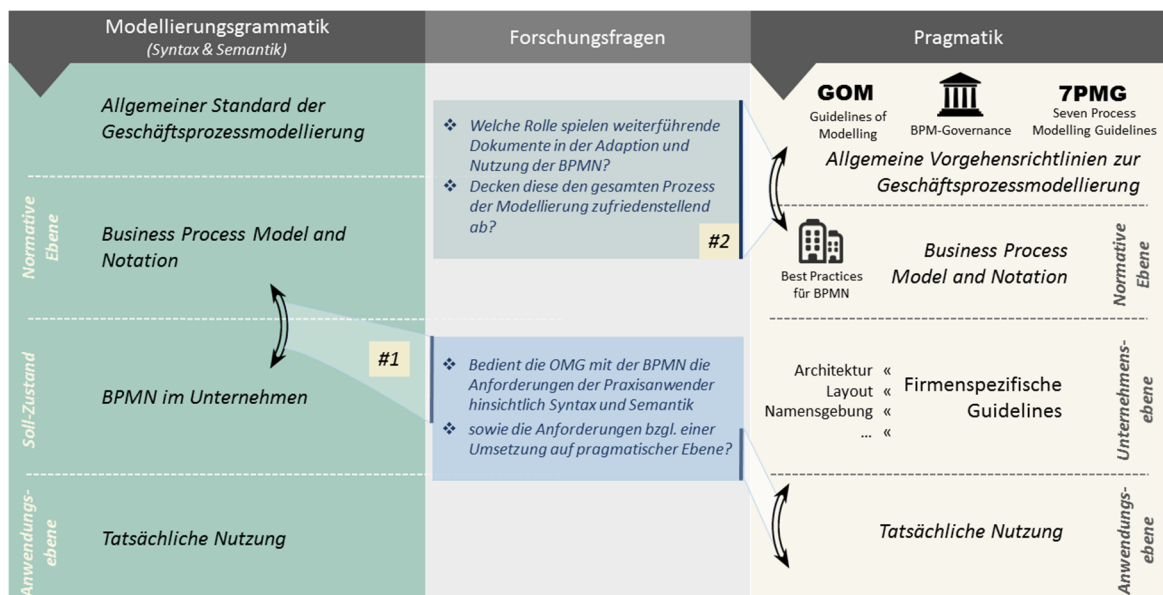
Die Evaluation der gestellten Forschungsfragen erfolgt auf der Basis des im Folgenden eingeführten Ebenen-Modells der Prozessmodellierung, das die BPMN unter dem Dach des BPM in vier Ebenen unterteilt. Das Modell dient der qualitativen Literaturanalyse als Framework zur Zuteilung ausgewählter Publikationen sowohl aus dem Umfeld der Modellierungsgrammatik als auch der Pragmatik der Modellierung. Eine graphische Auswertung der Analyse ermöglicht die Identifikation besonders relevanter Forschungsbeiträge. Aus der Literaturanalyse sind schließlich Rückschlüsse auf den in der Literatur empfohlenen Vorgang zur Prozessmodellierung möglich, welche bezüglich ihrer praktischen Umsetzung untersucht und mögliche Schwächen aufgedeckt werden.

## 2 Forschungsdesign

### 2.1 Ebenen-Modell der Prozessmodellierung

Die Literatur führt für die Wissenschaft der allgemeinen Zeichengesetze den Begriff der *Semiotik* ein (Lorenz 1995). Analog der Sprachentheorie sind auch in der Prozessmodellierung die drei miteinander in Referenz tretenden Ebenen der *Syntax*, *Semantik* und *Pragmatik* von höchster Bedeutung (Delfmann 2006). Dementsprechend grenzt auch das Ebenen-Modell der Prozessmodellierung die Anwendung der BPMN in einen grammatischen (Syntax und Semantik) und pragmatischen Aspekt ab. Neben dieser horizontalen Einteilung erscheint eine Unterteilung der Literatur in vier Ebenen absteigender Granularität mit folgender Begründung sinnvoll: Das BPM

vereint innerhalb seines geschlossenen Lebenszyklus ein Reihe von Management-Disziplinen, die auf der Basis inkrementeller und transformativer Verfahren die Effektivität von Geschäftsprozessen steigern (Snabe et al. 2008). Diese werden unter den *allgemeinen Standards zur Geschäftsprozessmodellierung* referenziert. Zur Erreichung seiner Ziele nutzt das BPM u.a. graphische Notationssprachen, die dem BPM direkt untergeordnet sind (*normativen Ebene*) (Weske 2012). In diesem Artikel beschränken wir uns auf die BPMN. Auf der normativen Ebene legen Spezifikationsdokumente ihre syntaktischen und semantischen Regeln fest. Die weiteren Ebenen (*Soll-Zustand*, *tatsächliche Nutzung*) zielen auf eine Validierung der BPMN auf die Anforderungen der Praxisanwender ab.



**Abbildung 1: Ebenen-Modell der Prozessmodellierung inklusive Verortung der Forschungsfragen**

Die Pragmatik bewegt sich auf einer nicht-wörtlichen Ebene und beurteilt im Kontext der Prozessmodellierung die Interpretation und Anwendung der Grammatik durch ein Individuum vor dem Hintergrund eines Modellzwecks (Delfmann 2006; Holten 2000). Wird der Zweck des Modells durch eine geeignete Nutzung von Syntax und Semantik erfüllt, gilt die pragmatische Bedeutung des Modells als erfüllt (Holten 2000). Dementsprechend können unter pragmatischen Hilfestellungen im Kontext des Ebenen-Modells jegliche Bemühungen (Dokumente, Schulungen etc.) referenziert werden, die eine Anwendung der Modellierungsgrammatik zweckorientiert anleiten. Dies ist im eingeführten Ebenen-Modell der Prozessmodellierung auf *allgemeiner*, *normativer* und *firmenspezifischer Ebene* der Pragmatik möglich. Die vierte Ebene evaluiert die Adaption der zur Verfügung gestellten Hilfestellungen seitens der Modellersteller, indem sie die Einhaltung der empfohlenen Vorgehensweisen prüft (*tatsächliche Anwendung*). Als Beispiele können abstrakte Handlungsempfehlungen, unternehmensspezifische Governance-Dokumente oder durch die Literatur bereitgestellte Handlungsrahmen und Konventionen für die Modellierung mit BPMN angeführt werden.

## 2.2 Ansatz zur Durchführung der Literaturanalyse

Die Ausführungen der Literaturanalyse orientieren sich an der durch vom Brocke et al. (2015) vorgeschlagenen Vorgehensweise, die zur Abarbeitung eines achtstufigen Analyseprozesses auffordert. Aus Gründen der Übersichtlichkeit wird an dieser Stelle lediglich auf die wichtigsten

Parameter der Literaturanalyse eingegangen (vgl. Abbildung 2). Da die Literaturanalyse als Grundlage für die Ausführungen nachfolgender Kapitel dient, findet ihr *Prozess* sequentiell, d. h. in sich geschlossen und chronologisch *vor* der Auswertung der eigentlichen Forschungsfragen statt (vom Brocke et al. 2015). Für die Recherche der *Quellen* wurden sowohl Indexierungsservices für Zitationen als auch bibliographische Datenbanken herangezogen. Neben der Literaturdatenbank *Google Scholar* wurden zusätzlich die Datenbanken der acht wichtigsten Journals im Forschungsfeld der Informationssysteme (IS), den *Senior Scholars' Basket of Journals* (Senior Scholars Consortium 2011), abgefragt. Insgesamt wurden in der repräsentativen Literaturrecherche rund 100 Publikationen ausgewertet und inhaltlich der Taxonomie des Ebenen-Modells zugeordnet. Dabei wurden die *Techniken* der *Schlüsselwortsuche*, *Rückwärtssuche* und *Vorwärtssuche* kombiniert eingesetzt.

1	Prozess	Sequentiell		Iterativ
2	Quelle	Indexierungsservices	Bibliographische Datenbanken	Publikationen
3	Abdeckung	Umfassend	Repräsentativ	Grundlegend
4	Techniken	Schlüsselwörter	Rückwärtssuche	Vorwärtssuche

**Abbildung 2: Suchbereich der Literaturanalyse**

Mit Suchparametern wie „BPM“, „Process Modeling“, „Syntax“ oder „Semantic“ begann die Schlüsselwortsuche mit einer Auswahl möglichst abstrakter Suchbegriffe, die durch die Kombination mit Begriffen wie „Usage“ oder „Quality“ in einen anwendungsbezogenen Kontext gesetzt wurden. Mit diesem Vorgehen sollte eine möglichst hohe Abdeckung relevanter Publikationen im Bereich des BPM, als auch der BPMN erreicht werden. Zudem wurde die Sortierreihenfolge „Relevanz“ gewählt, um besonders häufig zitierte Publikationen in den Vordergrund zu stellen. Durch die Schlüsselwortsuche konnten insgesamt 57 unterschiedliche Publikationen von 92 Autoren als relevant für die untersuchte Forschungsthematik eingestuft werden. Um für die anschließend durchgeführte Vorwärts- und Rückwärtsanalyse besonders aktive und angesehene Autoren auszuwählen, wurde für jeden der 92 Autoren ein Gewichtungsfaktor berechnet. Als Bewertungsgrundlage wurden die innerhalb der Schlüsselwortsuche aufgenommenen Publikationen sowie deren Zitationen zum Stichtag des 31. Oktober 2015 in der Literaturdatenbank *Google Scholar* herangezogen. Der Gewichtungsfaktor  $f$  bildete für jeden der 92 Autoren  $a$  den durchschnittlichen Re-Zitationswert  $z$  pro Publikation  $p$  (vgl. Gleichung (1) und (2)). Für eine Publikation  $p$  gilt  $p \in s$ , wobei  $s$  für die im vorherigen Absatz eingeführten Suchparameter steht. Der Endwert  $n$  leitet sich aus der Gesamtanzahl innerhalb der Suchwörter indexierter Publikationen ab.

$$f_a = \frac{\sum_{p=1}^n q_{pa} * z_p}{\sum_{p=1}^n q_{pa}}, \forall a. \quad (1)$$

$$\text{für } q_{pa} = \begin{cases} 1, & \text{falls Autor } a \text{ an Publikation } p \text{ mitgewirkt} \\ 0, & \text{sonst} \end{cases} \quad (2)$$

Für den Autor *Michael zur Muehlen* ergibt sich beispielsweise ein Gewichtungsfaktor  $f$  von 114 durchschnittlichen Zitationen pro Artikel. Hierfür wurden zunächst alle Artikel des Autors identifiziert, die als Ergebnis der Schlüsselwortsuche in die Literaturdatenbank aufgenommen wurden (3 aus 65). Für die drei identifizierten Artikel – zur Muehlen und Ho (2008), zur Muehlen und Recker (2008) sowie zur Muehlen et al. (2010) – wurden anschließend die Werte ihrer Re-Zitationen addiert. Die entstandene Summe (342) ergibt in der Division durch die Anzahl der Artikel (3) den Wert 114. Der durch Hirsch (2005) definierte h-Index schien in diesem Kontext als

ungeeignet, da durch diesen eine harmonische Angleichung vielzitiertes Artikel erfolgt und dieser Effekt für die Einstufung der Autoren als nicht wünschenswert erachtet wurde. Entsprechend des Beispiels wurde die Formel auf die aus der Schlüsselwortsuche verfügbare Literaturdatenbasis angewendet und die Autoren *Bandara, Becker, Indulska, Mendling, Recker, Rosemann, van der Aalst und zur Muehlen* als besonders relevant identifiziert. Mit einer Autorensuche wurde in der Literaturdatenbank Google Scholar anschließend für jeden Autor der meistzitierte Artikel mit einem Bezug zur Forschungsdomäne ausgewählt und eine Vorwärts- und Rückwärtsanalyse durchgeführt.

### 3 Literaturanalyse

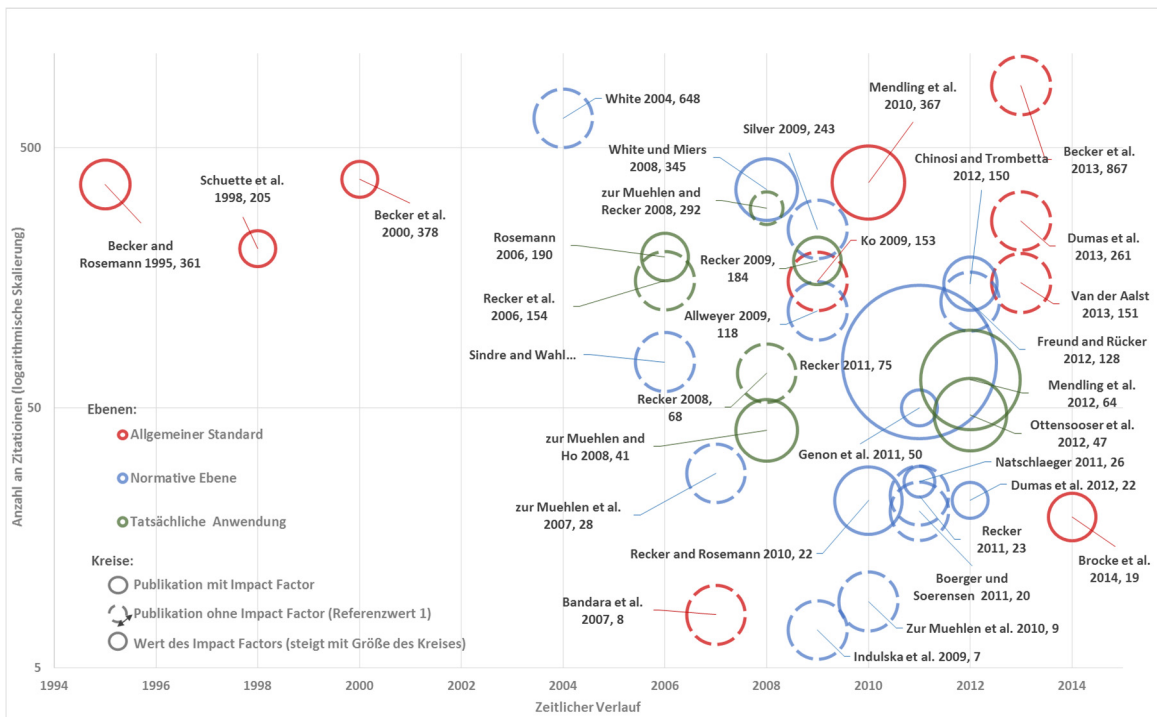
#### 3.1 Ergebnisse der Literaturanalyse

Die Ergebnisse der Literatursuche wurden sowohl aus modellierungsgrammatischer, als auch aus pragmatischer Sicht untersucht und in das Schema des Ebenen-Modells eingeordnet. Die Einteilung ermöglicht eine kumulative Auswertung der Publikationen und lässt erste Rückschlüsse über die Qualität der den einzelnen Ebenen zugeordneten Publikationen zu. Die Einordnung der evaluierten Literatur erfolgte anhand der (subjektiven) Einschätzung des Autors und wird an dieser Stelle aus Gründen der Übersichtlichkeit nicht näher erläutert.

Die Literatur der *Modellierungsgrammatik* zeichnet sich durch einen starken Bezug zur Syntaktik und Semantik der Sprache aus, ohne dabei auf pragmatische Ansätze einzugehen. Auf der Ebene des *allgemeinen Standards* werden Publikationen wie Ko et al. (2009) oder van der Aalst et al. (2003) referenziert, die das BPM und seine Methoden auf einer abstrakten und allgemein verständlichen Ebene darstellen. Auf der *normativen Ebene* finden sich vorrangig Spezifikationsdokumente der OMG, welche die BPMN in ihrer Syntax und Semantik definieren. Der *Soll-Zustand* beschreibt die Anforderungen der Praxisanwender bzw. der Unternehmen an die Modellierungsgrammatik eines Standards. Empirische Untersuchungen beurteilen auf dieser Ebene beispielsweise den Reifegrad der BPMN oder messen die Zufriedenheit der Anwender mit der Notation. Die *Anwendungsebene* bewertet schließlich die tatsächliche Nutzung des Standards und evaluiert die praxisbezogenen Modellierungserfahrungen mit dem Standard. Indulska et al. (2009b) stellen beispielsweise aktuelle Problemstellungen und zukünftige Herausforderungen in der Modellierung mit BPMN heraus.

Die *Pragmatik* der Modellierung verfügt über eine deutliche höhere Publikationsdichte. Die pragmatische Abhandlung der Notation wird in den einzelnen Ebenen vorrangig durch Handlungsanleitungen bzw. -empfehlungen, Modellierungskonventionen oder qualitätsfördernden Frameworks begleitet. Auf *oberster Ebene* setzen sich neben vereinzelt Einführungsdokumenten im Besonderen abstrakte Handlungsempfehlungen bezüglich des Vorgangs der Modellierung durch, die sich in ihren Ausführungen noch auf keine spezifische Notationssprache beschränken. Die *normative Ebene* legt ihren Fokus auf die BPMN als eingesetzte Modellierungstechnik und listet Publikationen, die sich mit dem Aufbau eines umfangreichen Verständnisses für die Notation, als auch mit der Qualitätsforschung und Einführung von Frameworks befassen. Auf dieser Ebene kann sich der Praxisanwender besonders durch Praxishandbücher ein umfassendes Bild über die Modellierung mit der BPMN bilden und wird sowohl im Modellerstellungsprozess, als auch bezüglich einer Evaluation des Modellzweckes weitestgehend angeleitet. Als ausschlaggebende Publikationen können in diesem Zusammenhang Freund und Rücker (2012), Silver (2009) sowie Allweyer (2009) genannt werden. Zudem wird durch Publikationen wie Recker et al. (2011), Indulska et al. (2009c) oder zur Muehlen et al. (2007) eine erhöhte Sensibilität für die qualitative

Beschaffenheit von Prozessmodellen geweckt. Auf der *Unternehmensebene* ist meist eine individualisierte und durch Konventionen eingeschränkte Version der BPMN zu finden, die innerhalb interner BPM-Governance-Strukturen verortet wurde (Khusidman 2010; Spanyi 2010). Diese fokussiert sich auf den unternehmensspezifischen Anwendungskontext der BPMN und lässt aus Sicht des Unternehmens unnötige Modelle, Konstrukte oder andere Spezifikationen außen vor (Lichka et al. 2011). Entsprechende Dokumente werden nur in seltenen Fällen extern veröffentlicht. Die Publikationen auf *Anwendungsebene* beschäftigten sich mit der Fragestellung, wie BPMN im unternehmensspezifischen Kontext wirklich genutzt wird. Im Kontrast zu gleicher in der Modellierungsgrammatik formulierten Fragestellung wird hier ein besonderer Fokus auf die Verfügbarkeit von Hilfestellungen gelegt, die dem Modellierer in Form von Schulungen, Handlungsanleitungen, Referenzmodellen oder ähnlichen Ressourcen zur Verfügung stehen. Entsprechende Publikationen ermöglichen innerhalb von empirischen Studien Einblicke in die Adaption der BPMN. zur Muehlen und Recker (2008) kamen beispielsweise zu dem Schluss, dass lediglich vier der 50 Modell-Konstrukte der BPMN in mehr als 50 % der untersuchten Modellierungsprojekte genutzt wurden.



**Abbildung 3: Graphische Darstellung der Ergebnisse der Literaturanalyse bzgl. der Pragmatik**

Abbildung 3 zeigt eine graphische Übersicht der ausgewerteten Literatur bzgl. der Pragmatik der Modellierung. Hierbei wird neben dem Publikationsjahr (*X-Achse*) und der Anzahl an Zitationen (*Y-Achse*) auch der Impact Factor des Journals, in dem der Artikel publiziert wurde, berücksichtigt (*Größe der Kreise*). Wurde kein Impact Factor gefunden, ist der Kreis mit dem Referenzwert 1 dimensioniert und mit einem nicht durchgezogenen Rahmen dargestellt. Die Y-Achse wurde aufgrund der hohen Publikationsdichte logarithmisch zur Basis 10 skaliert. Die als Kreise dargestellten Publikationen enthalten eine Beschriftung bezüglich des Verfassers, des Publikationsjahres und der Anzahl ihrer Zitationen durch Drittartikel. Ihre Farbe stellt die Kategorie dar, die den Publikationen ausgehend des Ebenen-Modells zugeweiht wurde. Betrachtet man den *zeitlichen Verlauf*, fällt eine erhöhte Publikationsdichte im Zeitraum ab 2006 auf. Vor dieser Zeit



wurden lediglich Forschungsarbeiten mit einführendem Charakter (Owen und Raj 2003) sowie allgemeine Grundsätze ordnungsgemäßer Modellierung (Becker et al. 1995; Schuette und Rotthowe 1998; Becker et al. 2000) als relevant eingestuft. Letztere haben keinen spezifischen Bezug zu der erstmals 2004 eingeführten BPMN, sondern behalten übergreifend für alle Modellierungstechniken des BPM ihre Gültigkeit. Die folgende Tabelle teilt die in Abbildung 3 referenzierten Publikationen den einzelnen Ebenen des eingeführten Modells zu.

<b>Allgemeiner Standard</b>	vom Brocke et al. (2014), Becker et al. (2013), Dumas et al. (2013), van der Aalst (2013), Mendling et al. (2010), Ko (2009), Bandara et al. (2007), Becker et al. (2000), Becker et al. (1995)
<b>Normative Ebene</b>	Freund und Rücker (2012), Chinosi und Trombetta (2012), Dumas et al. (2012), Börger und Sörensen (2011), Natschläger (2011), Genon et al. (2011), Recker (2011), Recker et al. (2011), zur Muehlen et al. (2010), Recker und Rosemann (2010), Indulska et al. (2009c), Allweyer (2009), Silver (2009), White und Miers (2008), zur Muehlen et al. (2007), Sindre und Wahl (2006), White (2004)
<b>Firmenspezifische Guidelines</b>	Vgl. Kapitel 3.1 war zu erwarten, dass hier keine Publikationen indiziert werden: Diese werden i.d.R. als vertraulich gehandelt
<b>Tatsächliche Anwendung</b>	Haisjackl et al. (2015)*, (Leopold et al. 2015)*, Mendling et al. (2012), Ottensooser et al. (2012), Recker (2010), zur Muehlen und Ho (2008), Rosemann (2006)
* Publikation ist nicht in graphischer Darstellung (Abbildung 3) zu finden, da diese zum Zeitpunkt der Erstellung noch nicht zitiert wurden.	

**Tabelle 1: Auflistung der Publikationen in Bezug auf die pragmatische Seite des Ebenen-Modells**

### 3.2 Kritik an BPMN: Fehlende pragmatische Hilfestellungen

Innerhalb empirischer Auswertungen stellt Recker (2008; 2010) fest, dass Modellierer trotz einer in der Theorie als ausgereift erscheinenden Modellierungsgrammatik weiterhin Nutzungsproblemen in der Anwendung der BPMN begegnen. Denn Abseits von Syntax und Semantik lässt die Spezifikation durch ihre Komplexität großen Raum für Mehrdeutigkeiten. Silver (2009) bestärkt, dass die Notation redundante Modellelemente und Muster enthält. Er stellt fest, dass die durch die Notation gegebenen unterschiedlichen Möglichkeiten zur Modellierung eines identischen Sachverhalts die Unsicherheiten des Anwenders im Umgang mit dem Standard steigern können. Zudem können entscheidende Fragestellungen bzgl. der Pragmatik der Notation, die sich im Lernprozess der Modellierung stellen, nicht beantwortet werden (Silver 2009). Aufgrund fehlender pragmatischer Hilfestellungen bleiben seitens der BPMN-Spezifikation Fragestellungen wie *„Ist ein Kunde Teil des Geschäftsprozesses oder agiert dieser außerhalb davon“* oder *„Wie kann ein Kommunikationsablauf innerhalb eines Geschäftsprozesses abgebildet werden, wenn eine Nachricht nur zwischen externen Entitäten verwendet werden darf?“* unbeantwortet. Die Folge ist ein Mangel an Klarheit und Effektivität, die sich in der Modellierung mit der BPMN zeigt (Silver 2009). Denn obwohl sich die aktuelle Spezifikation genau über die Wichtigkeit der Definition einer visuellen Notation bewusst ist, verwendet sie hierfür keine streng wissenschaftlich fundierte Basis, sondern verlässt sich auf den gesunden Menschenverstand, die Intuition des Modellierers und die Emulation gängiger Praktiken (Genon et al. 2011). Die Folge sind suboptimale Sprachinterpretationen sowie unklare Design-Entscheidungen, die einer wirksamen Kommunikation zwischen den Beteiligten im Weg stehen. Zur Verdeutlichung können zwei, in der Spezifikation genannte Aspekte angeführt werden: Einerseits erläutert die Spezifikation in Kapitel 2.1.3, der *Visual Appearance* (Optik), sinngemäß, dass *die Auswahl der Formen und Symbole, die innerhalb der Spezifikation als graphische Elemente bereitgestellt werden, als ein Schlüsselement von*

*BPMN anzusehen ist*. Gleichzeitig wird die Intention unterstrichen, dass mit der Notation die Definition einer Standard-Bildsprache verfolgt wird, die ausnahmslos durch *alle beteiligten Prozessmodellersteller erkennbar und zu verstehen* sein soll (Object Management Group 2013). Auf der anderen Seite steht die BPMN-Spezifikation der Bereitstellung einer eindeutigen pragmatischen Handlungsempfehlung, die essentiell für die Verwendung einer Sprache ist (Lorenz 1995; Delfmann 2006), gänzlich bzw. in zu großem Umfang nach. Es wird für die Erstellung von Prozessdiagrammen lediglich auf die Verwendung der in der Spezifikation bereitgestellten graphischen Elemente, Formen und Marker verwiesen (Object Management Group 2013). Sobald die Veröffentlichung einer Modellierungssprache erfolgt, wird diese in die Hand der Anwender gegeben. Während ab diesem Zeitpunkt der Einfluss der Sprachdesigner nachlässt, steigt die Wahrscheinlichkeit, dass die Anwender die Sprache entsprechend ihrer eigenen Intentionen und Zwecke einsetzen (zur Muehlen 2008). Um den erodierenden Effekt der Standardisierung zu vermeiden und somit die Bildung von Sprachdialekten bzw. einer vermehrt leichtgewichtigen Anwendung zu verhindern (Bjeković et al. 2014), sollte sowohl die Modellierungsgrammatik als auch im Besonderen die Pragmatik keine wichtigen Fragestellungen bezüglich der Anwendung offen lassen. Denn die Modellierung ist ein stark subjektiver, von den individuellen Eigenschaften eines Prozessmodellierers abhängiger Vorgang und erfordert eine pragmatische Beziehung zwischen Modell und Modellersteller (Holten 2000). Eine einheitliche und dem Zweck dienende Darstellung von Geschäftsprozessen ist nur dann möglich, wenn eine angemessene Beschreibung des pragmatischen Einsatzes vorausgesetzt werden kann (Silver 2009; van der Aalst et al. 2003). Die Literatur bietet in diesem Kontext zahlreiche Hilfestellungen an (vgl. Tabelle 1, *normative Ebene*), die das Prozessmodell meist auf qualitativer Ebene, d.h. als gegebenes und fertiggestelltes Artefakt betrachten (Claes et al. 2012). Auf diese Weise können beispielsweise Missstände in der Modellierung aufgedeckt, Modelle vereinfacht oder dessen Aussagekraft untersucht werden (zur Muehlen et al. 2010; Mendling et al. 2009).

### **3.3 Die kognitive Belastung im Modellerstellungsprozess**

Eine Anleitung des gesamten Modellerstellungsprozesses, inklusive der Phase der Anforderungsspezifikation an ein Prozessmodell noch *vor* der eigentlichen Modellierung, findet in der Literatur bisher kaum statt (Soffer et al. 2012). Mit der Identifikation, Bewertung und Validierung des zu modellierenden Sachverhalts steht die mentale Auffassungsgabe des Prozessmodellierers jedoch noch vor der eigentlichen Modellerstellung im Vordergrund (Weske 2012). In dieser Phase wird der Modellierer bereits mit der Konstruktion eines mentalen Modells und dessen anschließender Überführung in ein Prozessmodell konfrontiert. Innerhalb dieser Phase müssen nicht nur einfache zu erkennende Zuständigkeiten oder Aktivitäten identifiziert werden, sondern auch Routing-Entscheidungen getroffen und modelliert werden. Für die meisten Modellierer bewegt sich das Routing auf einer Komplexitätsebene, dessen Entscheidungswege oft weder eindeutig noch beobachtbar sind (Soffer et al. 2012). Um die Konstruktion des mentalen Modells erfolgreich zu gestalten, hat ein Modellierer *bereits vor der Umsetzung qualitativer Handlungsanweisungen* über ein grundlegendes Verständnis für die Domäne und dessen Verhalten zu verfügen (Soffer et al. 2012). Auch die anschließende Überführung des mentalen Modells in ein Prozessmodell, erfordert konzeptionelle und technische Fähigkeiten, die in Abhängigkeit der Ausdrucksmächtigkeit einer Sprache und den individuellen Fähigkeiten des Modellierers schnell überfordert werden können (Soffer et al. 2012).

## 4 Schlussbetrachtung

Während die qualitative Literaturanalyse bezüglich der Syntax und Semantik der BPMN keine ausschlaggebenden Mängel aufdeckt, wird seitens der Pragmatik bestätigt, dass Hilfestellungen hinsichtlich einer kontextabhängigen Anwendung der BPMN von offizieller Seite nicht gegeben werden. Aus der graphischen Darstellung der evaluierten Literatur kann abgeleitet werden, dass sich nach Einführung der BPMN im Jahr 2004 besonders ab 2006 ein erhöhtes Interesse an der Notation gezeigt hat. In Form einer angestiegenen Publikationsdichte kann der BPMN ebenfalls ein direkter Einfluss auf das Interesse für den allgemeinen Standard (BPM) zugesprochen werden: Die Anzahl der Publikationen haben sich auf dieser Ebene nach 2006 (bis 2012) im Gegensatz zu den Jahren 1994 bis 2000 deutlich vervielfacht (vgl. Abbildung 3). Zudem ermöglichen die Auswertungen der Literaturanalyse eine kritische Betrachtung der ausgegebenen Zielstellungen hinsichtlich der „Visual Appearance“ sowie der angestrebten umfassenden Verständlichkeit der Notation durch involvierte Prozessmodellersteller. Die Spezifikation kann diese ambitionierten Ziele ohne die Unterstützung fachbezogener Literatur nicht erreichen. Zwar stellt die OMG mit der BPMN eine ausreichende Basis zur syntaktisch aussagekräftigen und semantisch korrekten Abbildung von Geschäftsprozessen zur Verfügung, eine qualitativ hochwertige und auf den Kontext angepasste Modellierung ist jedoch nur unter Zuhilfenahme pragmatischer Hilfestellungen Dritter in Form akademischer oder unternehmensinterner Ergänzungen möglich. Hinzu kommen häufig zu beobachtende Mängel in der Modellierungsexpertise involvierter Prozessmodellierer (Pinggera et al. 2010; Suarez et al. 2011). Die Literatur versucht diesen Missstand durch entsprechende Fachbücher, Handlungsanleitungen oder Qualitätsframeworks auszugleichen, fokussiert sich mit ihren Bemühungen zur Verbesserung der Modellqualität jedoch meist auf die analytische Qualitätsprüfung. Die kognitive Belastung des Modellierers stellt dahingegen noch vor dem eigentlichen Modellerstellungsprozess und dem möglichen Einsatz prozessbegleitender Vorgehensrichtlinien bzw. Qualitätsframeworks einen wesentlichen Faktor für den Erfolg des Modellierungsvorganges dar (Soffer et al. 2012; Pinggera et al. 2013). Es folgt, dass pragmatische Hilfestellungen zwar in einem umfangreichen Ausmaß angeboten werden, diese aber nicht den gesamten, für die Qualität eines Prozessmodells entscheidenden Prozess abdecken. Aufgrund der angestellten Untersuchungen empfehlen die Autoren den derzeitigen Fokus pragmatischer Hilfestellungen auf die Modellerstellung und -evaluation zugunsten weiterführender Forschungen im Bereich der kognitiven Modellerstellung aufzuweichen und den Modellierer so bereits *vor* der eigentlichen Modellerstellung in der Konstruktion des mentalen Modells zu entlasten.

## 5 Literaturverzeichnis

- Allweyer T (2009) BPMN 2.0: Business Process Modeling Notation: Einführung in den Standard für die Geschäftsprozessmodellierung. Books on Demand, Norderstedt.
- Bandara W, Gable GG, Rosemann M (2007) Critical Success Factors of Business Process Modeling.
- Becker J, Kugeler M, Rosemann M (2013) Process Management: A Guide for the Design of Business Processes. Springer Science & Business Media.
- Becker J, Rosemann M, Schütte R (1995) Grundsätze ordnungsmäßiger Modellierung. Business and Information Systems Engineering 37 (5):435-445.
- Becker J, Rosemann M, von Uthmann C (2000) Guidelines of Business Process Modeling. In: van der Aalst W, Desel J, Oberweis A (Hrsg.) Business Process Management, vol 1806. Lecture Notes in Computer Science. Springer, Berlin, S. 30-49.

- Bjeković M, Proper H, Sottet J-S (2014) Embracing Pragmatics. In: Yu E, Dobbie G, Jarke M, Purao S (Hrsg.) *Conceptual Modeling*, vol 8824. *Lecture Notes in Computer Science*. Springer International Publishing, S. 431-444.
- Börger E, Sörensen O (2011) BPMN Core Modeling Concepts: Inheritance-Based Execution Semantics. In: Embley DW, Thalheim B (Hrsg.) *Handbook of Conceptual Modeling*. Springer Berlin, S. 287-332.
- Chinosi M, Trombetta A (2012) BPMN: An Introduction to the Standard. *Computer Standards & Interfaces* 34 (1):124-134.
- Claes J, Vanderfeesten I, Reijers HA, Pinggera J, Weidlich M, Zugal S, Fahland D, Weber B, Mendling J, Poels G (2012) Tying Process Model Quality to the Modeling Process: The Impact of Structuring, Movement, and Speed. In: *Business Process Management*. Springer, S. 33-48.
- Delfmann P (2006) *Adaptive Referenzmodellierung: Methodische Konzepte zur Konstruktion und Anwendung wiederverwendungsorientierter Informationsmodelle*. Dissertation, Münster.
- Dumas M, La Rosa M, Mendling J, Mäesalu R, Reijers H, Semenenko N (2012) Understanding Business Process Models: The Costs and Benefits of Structuredness. In: Ralyté J, Franch X, Brinkkemper S, Wrycza S (Hrsg.) *Advanced Information Systems Engineering*, vol 7328. *Lecture Notes in Computer Science*. Springer, Berlin, S. 31-46.
- Dumas M, La Rosa M, Mendling J, Reijers HA (2013) *Fundamentals of Business Process Management*. Springer, Heidelberg.
- Freund J, Rücker B (2012) *Praxishandbuch BPMN 2.0*, Bd. 3. Carl Hanser Verlag, München.
- Genon N, Heymans P, Amyot D (2011) Analysing the Cognitive Effectiveness of the BPMN 2.0 Visual Notation. In: Malloy B, Staab S, van den Brand M (Hrsg.) *Software Language Engineering*, vol 6563. *Lecture Notes in Computer Science*. Springer, Berlin, S. 377-396.
- Haisjackl C, Pinggera J, Soffer P, Zugal S, Lim SY, Weber B (2015) Identifying Quality Issues in BPMN Models: an Exploratory Study. In: *Enterprise, Business-Process and Information Systems Modeling*. *Lecture Notes in Business Information Processing*, vol 214. Springer, Berlin, S. 217-230.
- Hirsch JE (2005) An index to quantify an individual's scientific research output. *Proceedings of the National academy of Sciences of the United States of America* 102 (46):16569-16572.
- Holten R (2000) *Entwicklung einer Modellierungstechnik für Data-Warehouse-Fachkonzepte. Modellierung betrieblicher Informationssysteme* *Proceedings der MobIS-Fachtagung*. 7:3-21.
- Indulska M, Green P, Recker J, Rosemann M (2009a) Business Process Modeling: Perceived Benefits. In: *Conceptual Modeling-ER 2009*. Springer, S. 458-471.
- Indulska M, Recker J, Rosemann M, Green P Business Process Modeling: Current Issues and Future Challenges. *21st International Conference on Advanced Information Systems*, Amsterdam, 2009b. Springer, S. 501-514.
- Indulska M, zur Muehlen M, Recker J (2009c) *Measuring Method Complexity: The Case of the Business Process Modeling Notation*. *BPMcenter.org2009*.
- Khusidman V (2010) *BPM Governance Framework*. *BPTrends*.
- Ko RKL (2009) A Computer Scientist's Introductory Guide to Business Process Management (BPM). *Crossroads* 15 (4):11-18.
- Ko RKL, Lee SSG, Lee EW (2009) Business Process Management (BPM) Standards: A Survey. *Business Process Management Journal* 15 (5):744-791.
- Krogstie J, Sindre G, Jorgensen H (2006) Process Models Representing Knowledge for Action: A Revised Quality Framework. *European Journal of Information Systems* 15 (1):91-102.
- Leopold H, Mendling J, Gunther O (2015) *What we can Learn From Quality Issues of BPMN Models from Industry*.
- Lichka C, Boudinova D, Sommer J, Wittwer F (2011) BPMN Used by Business Professionals: An In-depth Reflection on BPM with BPMN by the Swis FOITT. In: Fischer L (Hrsg.) *BPMN 2.0*

- Handbook Second Edition: Methods, Concepts, Case Studies and Standards in Business Process Modeling Notation. Future strategies Inc., Lighthouse Point, FL, S. 243 - 252.
- Lorenz K (1995) Semiotik. In: Mittelstraß J (Hrsg.) Enzyklopädie Philosophie und Wissenschaftstheorie, vol 3. Stuttgart, S. 781-786.
- Mendling J, Recker J, Reijers HA (2009) Process Modeling Quality: A Framework and Research Agenda. BPM Center Report BPM-09-02, BPM Center Report org.
- Mendling J, Reijers HA, van der Aalst WM (2010) Seven Process Modeling Guidelines (7PMG). *Information and Software Technology* 52 (2):127-136.
- Mendling J, Strembeck M, Recker J (2012) Factors of Process Model Comprehension: Findings from a Series of Experiments. *Decision Support Systems* 53 (1):195-206.
- Muenstermann B, Eckhardt A, Weitzel T (2010) The Performance Impact of Business Process Standardization: An Empirical Evaluation of the Recruitment Process. *Business Process Management Journal* 16 (1):29-56.
- Natschläger C (2011) Towards a BPMN 2.0 Ontology. In: Dijkman R, Hofstetter J, Koehler J (Hrsg.) *Business Process Model and Notation*, vol 95. *Lecture Notes in Business Information Processing*. Springer Berlin Heidelberg, S. 1-15.
- Object Management Group (2013) *Business Process Model and Notation (BPMN). Version 2.0.2*. Object Management Group (OMG).
- Ottensooser A, Fekete A, Reijers HA, Mendling J, Menictas C (2012) Making Sense of Business Process Descriptions: An Experimental Comparison of Graphical and Textual Notations. *Journal of Systems and Software* 85 (3):596-606.
- Overhage PDS, Birkmeier DQ, Schlauderer S (2012) Quality Marks, Metrics, and Measurement Procedures for Business Process Models. *Business & Information Systems Engineering* 4 (5):229-246.
- Owen M, Raj J (2003) *BPMN and Business Process Management: Introduction to the New Business Process Modeling Standard*.
- Pinggera J, Furtner M, Martini M, Sachse P, Reiter K, Zugal S, Weber B (2013) Investigating the Process of Process Modeling with Eye Movement Analysis. In: La Rosa M, Soffer P (Hrsg.) *Business Process Management Workshops*, vol 132. *Lecture Notes in Business Information Processing*. Springer Berlin Heidelberg, S. 438-450.
- Pinggera J, Zugal S, Weber B, Fahland D, Weidlich M, Mendling J, Reijers H (2010) How the Structuring of Domain Knowledge Helps Casual Process Modelers. In: Parsons J, Saeki M, Shoval P, Woo C, Wand Y (Hrsg.) *Conceptual Modeling – ER 2010*, vol 6412. *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, S. 445-451.
- Recker J (2010) Opportunities and Constraints: The Current Struggle with BPMN. *Business Process Management Journal* 16 (1):181-201.
- Recker J (2011) Evaluations of Process Modeling Grammars: Ontological, Qualitative and Quantitative Analyses Using the Example of BPMN, Bd. 71. Springer Science & Business Media.
- Recker J, Rosemann M (2010) The Measurement of Perceived Ontological Deficiencies of Conceptual Modeling Grammars. *Data & Knowledge Engineering* 69 (5):516-532.
- Recker J, Rosemann M, Green PF, Indulska M (2011) Do Ontological Deficiencies in Modeling Grammars Matter? *Management Information Systems Quarterly* 35 (1):57-79.
- Recker JC (2008) BPMN Modeling: Who, Where, How and Why. *BPTrends* 5 (3):1-8.
- Rosemann M (2006) Potential Pitfalls of Process Modeling: Part A. *Business Process Management Journal* 12 (2):249-254.
- Schuetter R, Rotthowe T (1998) The Guidelines of Modeling: An Approach to Enhance the Quality in Information Models. In: Ling T-W, Ram S, Li Lee M (Hrsg.) *Conceptual Modeling – ER '98*, vol 1507. *Lecture Notes in Computer Science*. Springer, Berlin, S. 240-254.

- Senior Scholars Consortium (2011) Senior Scholar's Basket of Journals. <http://aisnet.org/?SeniorScholarBasket>. Accessed 6 Oktober 2015.
- Silver B (2009) BPMN Method and Style. Cody-Cassidy Press, Aptos, CA.
- Sindre G, Wahl T (2006) An Analytical Evaluation of BPMN Using a Semiotic Quality Framework. In: Advanced topics in database research, vol 5. Advanced Topics in Database Research. Siau, Keng, Columbia, MO, S. 94.
- Snabe JH, Rosenberg A, Miller C, Scavillo M (2008) Business Process Management: The SAP Roadmap. SAP Press.
- Soffer P, Kaner M, Wand Y (2012) Towards Understanding the Process of Process Modeling: Theoretical and Empirical Considerations. In: Daniel F, Barkaoui K, Dustdar S (Hrsg.) Business Process Management Workshops, vol 99. Lecture Notes in Business Information Processing. Springer Berlin Heidelberg, S. 357-369.
- Spanyi A (2010) Business Process Management Governance. In: vom Brocke J, Rosemann M (Hrsg.) Handbook on Business Process Management 2. International Handbooks on Information Systems. Springer Berlin Heidelberg, S. 223-238.
- Suarez GN, Freund J, Schrepfer M (2011) Best Practice Guidelines for BPMN 2.0. In: Fischer L (Hrsg.) BPMN 2.0 Handbook Second Edition: Methods, Concepts, Case Studies and Standards in Business Process Modeling Notation. Future strategies Inc., S. 203 - 216.
- van der Aalst WM (2013) Business Process Management: A Comprehensive Survey. ISRN Software Engineering 2013.
- van der Aalst WM, Ter Hofstede AH, Weske M (2003) Business Process Management: A Survey. In: Business Process Management. Springer, S. 1-12.
- vom Brocke J, Schmiedel T, Recker J, Trkman P, Mertens W, Viaene S (2014) Ten Principles of Good Business Process Management. Business Process Management Journal 20 (4):530-548.
- vom Brocke J, Simons A, Riemer K, Niehaves B, Plattfaut R, Cleven A (2015) Standing on the Shoulders of Giants: Challenges and Recommendations of Literature Search in Information Systems Research. Communications of the Association for Information Systems 37 (1):9.
- Weske M (2012) Business Process Management: Concepts, Languages, Architectures. Springer Science & Business Media.
- White SA (2004) Introduction to BPMN. IBM Cooperation 2.
- White SA, Miers D (2008) BPMN Modeling and Reference Guide: Understanding and Using BPMN : Develop Rigorous Yet Understandable Graphical Representations of Business Processes. Future Strategies Inc., Lighthouse Point, Fla.
- zur Muehlen M (2008) BPM Research and Education: Evidence-Based Research-What We Know, What We Don't and Why We Should Find Out. BPTrends 5:1-6.
- zur Muehlen M, Ho DT Service Process Innovation: A Case Study of BPMN in Practice. 41st Annual Hawaii International Conference on System Sciences, Waikoloa, Big Island, HI, 7-10 Jan. 2008 2008. S. 372-381.
- zur Muehlen M, Recker J (2008) How Much Language Is Enough? Theoretical and Practical Use of the Business Process Modeling Notation. In: Bellahsene Z, Léonard M (Hrsg.) Advanced Information Systems Engineering, vol 5074. Lecture Notes in Computer Science. Springer Berlin Heidelberg, S. 465-479.
- zur Muehlen M, Wisnosky DE, Kindrick J Primitives: Design Guidelines and Architecture for BPMN Models. 2010 Australasian Conference on Information Systems (ACIS), Brisbane, 2010. Australasian Computer Society.
- zur Muehlen Mz, Recker J, Indulska M Sometimes Less is More: Are Process Modeling Languages Overly Complex? EDOC Conference Workshop, 2007. EDOC '07. Eleventh International IEEE, Annapolis, MD, 15-16 Oct. 2007 2007. S. 197-204.

# Towards Taxonomy of Internet of Things-Product-Services

Christoph Klima<sup>1</sup>, David Heim<sup>1</sup>, and Axel Winkelmann<sup>1</sup>

<sup>1</sup> Universität Würzburg, Lehrstuhl für BWL und Wirtschaftsinformatik,  
{christoph.klima | david.heim | axel.winkelmann}@uni-wuerzburg.de

## Abstract

The "Internet of Things" (IoT) as a new, revolutionizing paradigm has become an important topic in Information Systems Research within the last years. Reviewing the latest scientific literature revealed that existing IoT-product-services were not focus of structuring and organization work in IS research. Applying taxonomies can bring this missing structure, as they help classifying objects of interest and thus support researchers and practitioners understand and analyze complex domains. In this paper, we develop a taxonomy for classifying IoT-product-services in order to generate insights regarding how value is generated in the IoT even across domain-, technology-, or business model-borders. To test the framework, we applied it to an empirical sample of 29 IoT-product-services.

## 1 Structuring and organizing Internet of Things-Product-Services

The "Internet of Things" (IoT) as a new, revolutionizing paradigm has become an important topic in Information Systems Research within the past years (French et al. 2015). Nevertheless, no consistent definition can be found in the academic literature (Kees et al. 2015). Several authors regard the IoT in general as a combination of physical and digital components to create new products and enable new business models (Wortmann and Flüchter 2015). In this paper, we follow a definition by Kees et al. (2015), which describes the IoT as "connectivity of physical objects (things), equipped with sensors and actuators, to the Internet via data communication technology, enabling interaction with and/or among these objects".

Numerous publications have analyzed the value creation process in the IoT and conclude that "Smart Objects" as hybrids, composed of elements from both the physical and digital worlds, are the base for developing digital services by adding sensing and connectivity (Fleisch et al. 2014; Turber and Smiela 2014).

Today, there are more and more previously non-digital products (e.g. bikes, watches, energy meters) that are equipped with connectivity, which leads to numerous areas of applications for the IoT like Energy, Healthcare or Home, changing the nature of business models in these industries (French et al. 2015). Until 2020 there will be 26 billion "Smart Things" installed, generating a market size of

about \$7.1 trillion (Kees et al. 2015; Wortmann and Flüchter 2015). Nevertheless, reviewing the latest scientific literature revealed that existing IoT-product-services were not focus to structuring and organization work. Applying taxonomies can bring this missing structure (Nickerson et al. 2013; Katsma et al. 2010).

Taxonomies in IS research help classifying objects of interest and thus support researchers and practitioners understand and analyze complex domains (Nickerson et al. 2013). According to Nickerson et al. (2013), a taxonomy can be defined as a set of dimensions, each consisting of different, mutually exclusive and collectively exhaustive characteristics such that each object under consideration has one and only one characteristic for each dimension. In different publications, taxonomies are furthermore regarded both as the process of building a framework of dimensions and characteristics on the one hand and as using it for categorizing objects of interest on the other hand. In this paper, we regard a taxonomy as the result of a categorization approach (e.g. Haas et al. 2014).

The goals of developing a taxonomy in a certain area of research can be diverse and have at least an influence on its development (van der Walk and Axelsson 2015). In this paper, we pursue the goal of gaining knowledge regarding the development of IoT-product-services by structuring existing IoT-product-services. We want to figure out, how value is generated in the IoT even across domain-, technology-, or business model-borders. Hence, our research question is:

*(RQ) How is a taxonomy for IoT-product-services to be designed?*

This research question implies that in this paper we are not going to develop a completed taxonomy for IoT-product-services in the sense of the result of the classification process, rather than a classification framework for achieving this goal. Hence, this paper is "work in progress".

To answer the research question, the remainder of this paper is structured as follows: section 2 gives an overview of related work in the field of taxonomy building in IS research. Sections 3 and 4 introduce the research method and how we applied it to develop the taxonomy. Finally, in sections 5 an outlook and further research activities are discussed.

## **2 Related work**

Building taxonomies is widely used in IS research. There are numerous fields of application, where taxonomies support in structuring and understanding different topics. To give a brief overview, the following section presents exemplary scientific publications to demonstrate the wide range of taxonomy application. These publications were selected systematically as a representative set for this work in progress paper.

One area of application where numerous publications on taxonomy building can be found is mobile business (e.g. Nickerson et al. 2013; Heinonen and Pura 2008; Nickerson et al. 2007). Nonetheless, there are many other fields in IS that have been subject to taxonomy research, especially in recent years, like e.g. taxonomies for E-Government Services (Nusir and Bell 2013), Cybercrimes (Land et al. 2013), Crowdfunding (Haas et al. 2014), Digital Music Services (Katsma and Spil 2008) or Cryptocurrencies (Glaser and Bezenberger 2015).

Reviewing the scientific literature revealed that taxonomy research is widely used in different areas of application, but that a solid and applicable taxonomy for IoT-product-services does not exist.



### 3 Research method

After reviewing the scientific literature, we identified an often cited approach for creating taxonomies in IS research, developed by Nickerson et al. (2013). We applied that research process, which is aligned with the design science paradigm by Hevner et al. (2004) ("design as a search process") and consists of seven steps. It combines an empirical and a conceptual approach in an iterative research procedure (cf. figure 1). Applying this approach guides for developing useful taxonomies "but not necessarily 'best' or 'correct' ones, as these cannot be defined" (Nickerson et al. 2013).

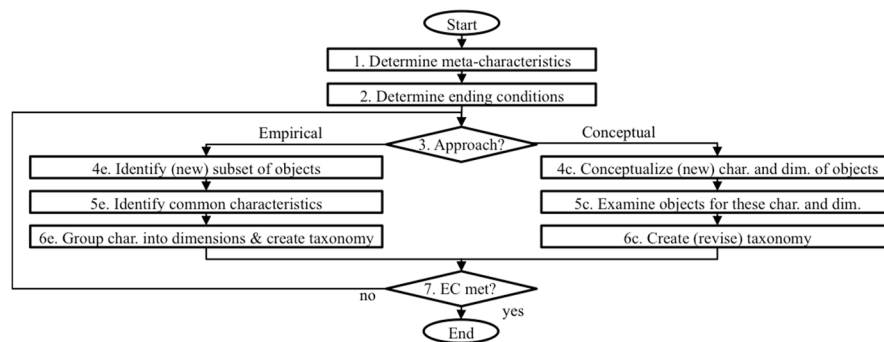


Figure 1: Research process by Nickerson et al. 2013

The applied research approach will be described in the next section, referring to our conducted research activities for taxonomy development in the domain of IoT-product-services.

### 4 Taxonomy development

(1) According to Nickerson et al. (2013), meta-characteristics should be defined before determining dimensions and characteristics in the taxonomy. Those meta-characteristic should be based on the purpose of the taxonomy. We followed the product-service-logic, presented by Fleisch et al. (2014) and Wortmann and Flüchter (2015) to derive our meta-categories "Smart Object" and "Service".

(2) Nickerson et al. (2013) furthermore defined a set of ending conditions for determining when to terminate the iterative research procedure. *Subjective* ending conditions are, that the taxonomy must satisfy the definition of a taxonomy and that it is concise, robust, comprehensive, extendible and explanatory. *Objective* conditions can be diverse. We defined the following ones, as also stated by Nickerson et al. (2013): at least one object is classified under every characteristics of every dimension; no new dimensions or characteristics were added in the last iteration; every dimension is unique and not repeated (i.e., there is no dimension duplication); every characteristic is unique within its dimension (i.e., there is no characteristic duplication within a dimension).

(3) In the next step, choosing the approach for the current iteration of the taxonomy building process "depends on the availability of data about objects under study and the knowledge of the researcher about the domain of interest" (Nickerson et al. 2013). It is useful to begin with a conceptual approach, if little data are available but the researcher has significant understanding of the domain. In our case, we started with a set of 13 IoT-product-services, which we did not regard as sufficient for choosing an empirical approach and thus decided to do a conceptual attempt first.

(4c) In the conceptual approach, the researcher begins by conceptualizing the dimensions of the taxonomy without examining actual objects. "This process is based on the researcher's notions about how objects are similar and how they are dissimilar" (Nickerson et al. 2013). Thus, we performed a review of the classification literature for our chosen meta-categories, identified a set of research papers within each meta-category and derived 19 generally applicable dimensions in the first iteration of the research procedure.

(5c) In the following step of the conceptual approach according to Nickerson et al. (2013), the researcher examines objects for these newly developed dimensions and characteristics and figures out, if there are objects that have each of the characteristics in each dimension. We used a test set of 13 IoT-product-services and were able to classify each of them according to all dimensions and characteristics. Hence, we were able to create a first version of the taxonomy (6c).

(7) After that, "the researcher asks if the ending conditions have been met with the current version of the taxonomy" (Nickerson et al. 2013). In our case, the taxonomy was not already concise, comprehensive and explanatory, as it contained too many dimensions and characteristics. Therefore, we decided to conduct a second iteration and chose an empirical approach.

(4e) In this step, an either random, systematic or convenience sample of objects is identified. In our case, we selected a random sample of 21 IoT-product-services that we were familiar with.

(5e) The next step of the empirical approach includes identifying common characteristics of the identified objects. Once a set of characteristics has been identified, they can be grouped e.g. informally using a manual process (Nickerson et al. 2013). We analyzed the subset of 21 IoT-product-services in detail in a group of three researchers in a workshop and developed 11 additional characteristics. Subsequently, we grouped them into four additional dimensions and thus created a second version of the taxonomy (6e).

As our taxonomy now included 23 dimensions, we did not regard it as meeting the subjective ending conditions (concise, robust, comprehensive, extendible) and decided to make a third iteration, choosing again a conceptual approach.

This time, we focused a lot more on the uniqueness of the dimensions and on their relevance according to our goal, which is gaining insights on existing IoT-product-services for developing new IoT-product-services, rather than on every generally applicable characteristic. Thus, we were able to remove 11 dimensions and reduced the amount to 12.

Finally, we conducted a fourth iteration, using an empirical approach again. We enlarged the sample to 29 IoT-product-services and analyzed again their common characteristics. We could not add another dimension to the taxonomy, but remove two more characteristics. During this iteration, at least one object was classified under every characteristics of every dimension, no new dimensions or characteristics were added, every dimension is unique and not repeated and every characteristic is unique within its dimension. Hence, the objective ending conditions could be met. As the taxonomy finally contains 12 dimensions and a number of characteristics from between two and five, it was regarded as concise, robust, comprehensive and extendible. Hence, we decided to finish the development process. Table 1 shows the framework as output of the applied research procedure.

Conceptually derived dimensions and characteristics are aligned with scientific literature for the meta-categories "Smart Object" (e.g. Hernández & Reiff-Marganiec 2014; Kortuem et al. 2010; Meyer et al. 2009) and "Service" (e.g. van der Walk and Axelsson 2015; Sorathia et al. 2010; Lovelock 1983).

Smart Object	Capability Levels	Networked		Enhanced		Aware		IoT Complete
	Data Analytics	Descriptive		Predictive		Exploratory		Prescriptive
	Data Source	User			User's property		Environment	
	Communication	Service to user				Bidirectional		
Service	Customization	Fully				Hardly		
	Breadth	Single Service				Package of services		
	Timing of Benefits	During Service delivery				After Service delivery		
	Domain	Home		Energy		Health		City
	Duration	On Demand				Continuous		
	Object affected	User				User's property		
	Market	Mass				Niche		
	Motivation	Economic/Rationality		Status/Recognition		Comfort	Security	Ecologic

**Table 1: Classification Framework for building taxonomy of IoT-product-services**

- The Smart Object meta-category consists of the dimensions "Capability Level" (which describes a Smart Object's functionality from only networked to fully goal oriented in four steps), "Data Analytics" (which describes the type and level of analytics, applied to IoT data), "Data source" (describes, whether IoT data is collected from the user, his or her property or from the environment" and "Communication" (describes the direction of communication and can either be bidirectional or unidirectional).
- The meta-category "Service" is composed of "Customization" (extent to which an IoT-product-service is customizable), "Breadth" (single service or package of services), "Benefit timing" (distinguishes IoT-product-services regarding the fact, at which time a user benefits from using it), "Domain" (scope of the IoT-product-service), "Duration" (defines if the IoT-product-service is only used on demand or continuous), "Object affected" (provides information about who or what is the recipient of the IoT-product-service), "Market" (describes, whether an IoT-product-service is provided to a niche market (e.g. sportsman, pet owners, homeowners) or to a mass market) and "Motivation" (describes the main consumer motivation, respectively value proposition, behind the usage of the IoT-product-service).

After developing the taxonomy, we applied it to 29 random IoT-product-services (cf. table 2), which proves that both objective and subjective ending conditions have been met and a usable taxonomy has been developed in four iterations, following strictly the research procedure stated by Nickerson et al. (2013).

	Service												Smart Object																							
	Motivation	Breadth	Market	Customization	Object	Benefit	Duration	Domain	Capability Level	Analytics	Communication	Data Source																								
<b>IoT-product-services</b>	Economic	Status	Comfort	Security	Ecological	Single Service	Package of Services	Mass	Niche	Customizable	Not Customizable	User	User's Property	During Service Delivery	After Service Delivery	On Demand	Continuous	Home	Energy	Health/Sport	Everyday Life	Networked	Enhanced	Aware	Complete	Descriptive Analytics	Predictive Analytics	Exploratory Analytics	Prescriptive Analytics	Service to User	Bidirectional	User	User's Property	Environment		
Adheretech Pill Bottles	1																																			
Adidas MiCoach smart ball	1																																			
Air Quality Egg																																				
AliveCor: Heart Monitor																																				
Aros Smart Window Air Conditioner																																				
Atlas Fitness Tracker	1																																			
August: Smart lock																																				
Awair: air monitor																																				
Beep Dial	1																																			
Belkin WeMo: Insight Switch																																				
Bionym Nymi	1																																			
BloomSky: Smart WeatherCam																																				
Breeze: Wireless breathalyzer																																				
British Gas Hive Active Heating 2																																				
Canary																																				
Coin 2.0	1																																			
ConnectSense Smart Outlet																																				
CycleAT: Tire sensors	1																																			
Deeper	1																																			
Duet: Smart bluetooth tag																																				
Flic Button																																				
Gecko: Bluetooth smart tag																																				
Goji Smart Lock																																				
Honeywell: Wifi smart thermostat	1																																			
Honeywell: Wifi smart thermostat																																				
Iro Smart sprinkler controller																																				
iSamrt Alarm																																				
Kinsa: Smart thermometer	1																																			

Table 2: Classification Framework applied to n=29 random IoT-product-services

### 5 Conclusion and Outlook

The Internet of Things (IoT) as important topic in Information Systems research is not already structured by applying taxonomies. In this work in progress paper we develop a taxonomy for classifying existing IoT-product-services. Our goal is to generate insights regarding how value is generated in the IoT even across domain-, technology-, or business model-borders. In a group of three researchers, following an approach by Nickerson et al. (2013), we were able to develop a taxonomy, consisting of 12 dimensions in four iterations. We successfully applied it to 29 IoT-product-services and met all of our subjective and objective ending conditions (goals).

As this paper is work in progress, the next steps will be to extend the empirical sample of IoT-product-services to a three-digit number and afterwards apply a cluster analysis as a statistical approach to develop a completed taxonomy for IoT-product-services in the sense of the result of the classification process.

### 6 Literature

Di Valentin C, Werth D, Loos P (2014) Towards a Business Model Framework for E-learning Companies. In: Proceedings of the AMCIS 2014. Savannah

Fleisch E, Weinberger M, Wortmann F (2014) Business Models and the Internet of Things. Bosch Internet of Things & Service Lab

French AM, Shim JP, Louis RS, Stevens B, Clark J (2015) Sailing the Seven Seas, a Blue Ocean of the Internet of Things. In: Proceedings of the AMCIS 2015. Puerto Rico

- Glaser F, Bezenberger L (2015) Beyond Cryptocurrencies - A Taxonomy of Decentralized Consensus Systems. In: Proceedings of the ECIS 2015. Münster
- Haas P, Blohm I, Leimeister JM (2014) An Empirical Taxonomy of Crowdfunding Intermediaries. In: Proceedings of the ICIS 2014. Auckland
- Heinonen K, Pura M (2008) Classifying Mobile Services. In: Proceedings of All Sprouts Content.
- Hernández M, Reiff-Marganiec S (2014) Classifying Smart Objects using Capabilities. In: Proceedings of the International Conference on Smart Computing 2014. Hong Kong
- Hevner AR, March ST, Park J, Ram S (2004) Design Science in Information Systems Research. *MIS Quarterly* 28(1):75-105
- Katsma C, Spil T (2008) A taxonomy of digital music services. In: Proceedings of the AMCIS 2008. Lima
- Kees A, Oberlaender AM, Roeglinger M, Rosemann M (2015) Understanding the Internet of Things: A Conceptualisation of Business-to-Thing (B2T) Interactions. In: Proceedings of the ECIS 2015. Münster
- Land L, Smith S, Pang V (2013) Building a Taxonomy for Cybercrimes. In: Proceedings of the PACIS 2013. Jeju Island
- Lovelock CH (1983) Classifying services to gain strategic marketing insights. *Journal of Marketing* 47:9-20
- Meyer GG, Främling K, Holmström J (2009) Intelligent Products: A survey. *Computers In Industry* 60(3):137-148
- Nickerson RC, Varshney U, Muntermann J, Isaac H (2007) Towards a Taxonomy of Mobile-product-services. In: Proceedings of the AMCIS 2007. Keystone
- Nickerson RC, Varshney U, Muntermann J (2013) A method for taxonomy development and its-product-service in information systems. *European Journal of Information Systems* (2013) 22:336–359
- Nusir M, Bell D (2013) Systematic Literature Review: Taxonomy Of Services In E-Government. In: Proceedings of UK Academy for Information Systems Conference 2013. Oxford
- Osterwalder A, Pigneur Y, Tucci CL (2005) Clarifying Business Models - Origins, Present, and Future of the Concept. *Communications of the Association for Information Systems* 16:1-25
- Sorathia V, Ferreira Pires L, van Sinderen M (2010) An Analysis of Service Ontologies. *Pacific Asia Journal of the Association for Information Systems* 2(1):17-46
- Kortuem G, Fitton D and V Sundramoorthy (2010) Smart Objects as Building Blocks for the Internet of Things. *IEEE Internet Computing* 14(1):44-51
- Turber S, Smiela C (2014) A Business Model Type for the Internet of Things. In: Proceedings of the ECIS 2014. Tel Aviv
- van der Valk W, Axelsson B (2015) Towards a managerially useful approach to classifying services. *Journal of Purchasing & Supply Management* 21(2):113–124
- Wortmann F, Flüchter K (2015) Internet of Things - Technology and Value Added. *Business Informations Systems Engineering* 57(3):221–224



# Effekte mobiler Audience Response Systeme in der universitären Lehre am Beispiel von eKARuS

Matti Klinsmann<sup>1</sup>, Philipp Melzer<sup>2</sup>, Bernd Schneider<sup>2</sup> und Mareike Schoop<sup>2</sup>

<sup>1</sup> Universität Stuttgart, matti@klinsmann.de

<sup>2</sup> Universität Hohenheim, Institut für Interorganisational Management & Performance, {philipp.melzer | bernd.schneider | schoop}@wi1.uni-hohenheim.de

## Abstract

Audience Response Systeme (ARS) können die Teilnahme, Interaktivität, Transparenz sowie die Motivation und schlussendlich den Lernerfolg in der universitären Lehre erhöhen. Sie werden seit den 90er Jahren zumeist verbunden mit proprietärer Hardware eingesetzt. Mit dem Einzug mobiler Endgeräte in Alltag und Lehre wurde der Zugang zu ARS weiter vereinfacht und neue Anwendungsszenarien ermöglicht. Der vorliegende Beitrag untersucht die Effekte eines mobilen ARS in einer realen Vorlesung an einer deutschen Universität. Hierbei wird die reine Interaktionshäufigkeit sowie der Lernerfolg anhand von Erinnerung und Verständnis der Vorlesungsinhalte quantitativ untersucht. Während die Ergebnisse hinsichtlich eines gesteigerten Interaktionsaufkommens bisherige Studien mit herkömmlichen ARS bestätigen, sind die Ergebnisse hinsichtlich des Lernerfolgs gemischt.

## 1 Audience Response Systeme im Wandel der Zeit

Audience Response Systeme (ARS) werden seit ihrer Einführung in den 60er Jahren (Froehlich 1963) und speziell mit dem Einzug der Infrarottechnologie im Jahr 1999 (Kay und LeSage 2009b) zunehmend in der universitären Lehre eingesetzt.

Neben der Anregung zur verstärkten Teilnahme an der Lehrveranstaltung (Draper und Brown 2004) erhoffen sich Lehrende erhöhte Interaktivität, Spaß und Transparenz über den Lernerfolg durch das sofortige Feedback. Diese Effekte wurden in zahlreichen Experimenten und Studien, vor allem in den frühen 2000er Jahren, untersucht und bestätigt (Fies und Marshall 2006; Graham et al. 2007; Kay und LeSage 2009a; Boyle und Nicol 2003; Latessa und Mouw 2005).

Während die Publikation neuer Studien gegen Ende der 2000er Jahre stetig abnahm, trug das Festhalten am traditionellen Funktionsprinzip mit proprietärer Hardware aktuellen Trends zunehmend weniger Rechnung. Mit dem Einzug mobiler Technologien (Naismith et al. 2004) und der seit 2010 stark ansteigenden Verbreitung von Smartphones (Gartner 2015), entwickelt sich auch die Art zu lernen und zu lehren weiter: Beispielsweise wurde in den Untersuchungen von Boyle

und Nicol (2003) angemerkt, dass die Informationsmenge, der der Mensch täglich ausgesetzt ist, zu kürzeren Aufmerksamkeitsspannen und einer Tendenz zum Multitasking führen kann.

Dies ist mit traditionellen, von langen Frontalvorträgen geprägten, Vorlesungsszenarien nur schwer vereinbar, was den Einsatz von ARS attraktiver macht. Zudem zeigt die zunehmende Integration mobiler Geräte in die Lehre im Rahmen des *Mobile Learning* auch an anderen Stellen Potentiale zur Verbesserung der Lernqualität (Cochrane 2010). Im Rahmen dieser Trends werden die Möglichkeiten drahtloser Verbindungen und mobiler Technologien zunehmend ausgeschöpft (Helmerich und Scherer 2007). Das webbasierte ARS *eKARuS*, welches am Lehrstuhl für Wirtschaftsinformatik 1 der Universität Hohenheim entwickelt wurde, vereint mehrere dieser Weiterentwicklungsansätze und dient bei der folgenden Untersuchung als Versuchsartefakt.

Auf Grundlage der Ergebnisse früherer Studien hat diese zum Ziel die Effekte von ARS zu identifizieren, welche im Zuge oben genannter Entwicklungen noch immer Gültigkeit besitzen, vor allem in Bezug auf den Lernerfolg. Hierzu wurde *eKARuS* im Rahmen eines kurzen zeitlichen Ausschnitts einer universitären Vorlesung im Hinblick auf Interaktionshäufigkeit und Lernerfolg explorativ evaluiert.

## 2 Effekte von Audience Response Systemen

Im ersten Schritt soll ein komprimiertes, zielgerichtetes Literaturreview die Einordnung und lehrtheoretische Herkunft der ARS beleuchten, um dann die Ergebnisse früherer Untersuchungen zusammengefasst darzustellen. Dabei wird der Fokus auf das Forschungsgebiet prägende Literatur gelegt, also Quellen, die entweder signifikante neue Aussagen zu betrachteten Effekten oder ergänzende Erkenntnisse enthalten, die diese Effekte stützen.

ARS sind dem Bereich *E-Learning* zuzuordnen und erben die durchgängige Charakteristik eines Begriffs, der zahlreiche durchaus heterogene Bereiche unter sich vereint: E-Learning wird als die Verwendung von verschiedenen technologischen, webunterstützten Werkzeugen verstanden, die für Lehrzwecke nutzbar sind (Nichols 2003). Der Mehrwert liegt dabei in der Unterstützung von Kommunikation und Denkprozessen um Wissen zu festigen. E-Learning hat damit nicht nur eine Unterstützungsfunktion in der Lehre, sondern birgt im Gegenteil auch erhebliches Transformationspotenzial und ermöglicht neue bzw. alternative pädagogische Ansätze (Garrison 2011).

Im Zusammenhang mit ARS sind zwei Strömungen von besonderer Bedeutung: die behavioristisch geprägte Theorie des aktiven Lernens und der soziale Konstruktivismus. Erstere besagt, dass Menschen durch aktives Lösen von Problemen lernen (Anzai und Simon 1979) und dies durch Strategien für die Vorlesung unterstützt werden soll (Bonwell und Eison 1991).

Letztgenannter besagt nach Vygotskiï und Cole (1978, S. 90), dass Lernprozesse erst auf sozialer Ebene, dann auf interner Ebene stattfinden. Ein großer Teil des individuellen Wissens stammt demnach aus Interaktionen mit anderen Menschen (Dufresne et al. 1996). Diese Erkenntnis wurde in weiteren Experimenten bestätigt (Doise et al. 1975; Springer et al. 1999) und lieferte zusammen mit den Implikationen aus der Theorie des aktiven Lernens eine klare Tendenz weg von traditionellen, einseitigen Vorlesungsszenarien hin zu interaktiveren Veranstaltungen (Draper und Brown 2004). Diese Entwicklung wurde auch auf Seiten der Lehrenden begrüßt und deren Auswirkung auf die Effektivität bestätigt (Butler 1992; Bonwell und Eison 1991).

Ein Beispiel für die konkrete Umsetzung zeigt das Konzept *Flipped Classroom*, bei dem der Vorlesung vorgelagert, individuelle computerbasierte Wissensvermittlung betrieben und dieses



Wissen in einem interaktiven Gruppenszenario während der Vorlesung angewendet bzw. erklärt wird (Bishop und Verleger 2013). Ähnlich verfährt auch der Ansatz der *Peer Instruction*, bei welchem die Studierenden während der Vorlesung Kernkonzepte aktiv auf Problemstellungen anwenden und diese vor der Zuhörerschaft erklären (Crouch und Mazur 2001).

Ebenso stellen ARS ein Mittel zur Erhöhung der Interaktivität dar. Die Lernenden werden mit einem mobilen Abstimmgerät ausgestattet (Burnstein und Lederman 2003) und stimmen damit über vom Lehrenden meist per Beamer oder Projektor gezeigte Umfragen ab. Die Ergebnisse dieser Abstimmung werden daraufhin unmittelbar für alle und meist in Diagrammform visualisiert (Kay und LeSage 2009a; Graham et al. 2007; Fies und Marshall 2006).

In vielen empirischen Untersuchungen zeigten sich die positiven Effekte eines solchen Systems. Folgende zentrale Erkenntnisse konnten mehrfach bestätigt werden und sollen daher als Basis für die Hypothesen des Experiments dienen.

ARS fördern die Teilnahme und aktive Mitarbeit der Studierenden an der Vorlesung (Caldwell 2007; Efstathiou und Bailey 2012; Latessa und Mouw 2005). Es findet demnach ein Abbau der Hemmschwelle zur Interaktionsinitiierung statt und das reine Interaktionsaufkommen steigt, was zu folgenden Hypothesen führt:

**H1:** ARS reduzieren die Hemmschwelle für Interaktionen seitens der Studierenden in Vorlesungen.

**H2:** ARS erhöhen das Interaktionsaufkommen zwischen Studierenden und dem Dozenten.

Außerdem wird ein höherer Spaßfaktor sowie höhere Motivation beobachtet, auch bei weniger interessanten Themen (Dufresne et al. 1996; Draper et al. 2002; Schmidt 2011). Die Studierenden beschäftigen sich folglich verstärkt mit den Inhalten der Vorlesung und es lässt sich entsprechend in einer Hypothese formulieren:

**H3:** ARS steigern die Auseinandersetzung der Studierenden mit Inhalten der Vorlesungen.

Als Konsequenz wird ein besseres Behalten von Vorlesungsinhalten, das korrekte Verständnis derselben (Boyle und Nicol 2003; Tregonning et al. 2012) und damit schließlich ein höherer Lernerfolg beobachtet (Scheele et al. 2004; Schackow et al. 2004; Wenz et al. 2013). Zusammengefasst lassen sich aus diesen Erkenntnissen folgende Hypothesen ableiten:

**H4:** ARS verbessern das Verständnis der Vorlesungsinhalte bei den Studierenden.

**H5:** ARS wirken sich positiv auf das Behalten der Vorlesungsinhalte bei den Studierenden aus.

Neben positiven Aspekten der Nutzung von ARS werden allerdings auch negative Punkte in den Untersuchungen angemerkt: So werden das Ablenkungspotenzial (Scheele et al. 2004), die Verwendung der Technologie um ihrer selbst willen (Draper und Brown 2004) und die oft auftretenden technischen Schwierigkeiten (Graham et al. 2007) kritisch betrachtet.

Letztere treten vor allem in Verbindung mit der proprietären Hardware, also mit Abstimmgeräten und Empfängereinheit, auf (Draper und Brown 2004). Diese führt zusätzlich zu signifikanten organisatorischen Problemen finanzieller und logistischer Natur (Fies und Marshall 2006). Die Ausstattung von Studierenden und/oder Hörsälen mit der entsprechenden Technik sowie deren Pflege erfordert hohe Investitionen (Dufresne et al. 1996; Latessa und Mouw 2005; Cain und Robinson 2008). Gleichzeitig stehen die Universitäten vor dem Problem, die Technik meist nur an

einem Ort einsetzen zu können, da die Hardware in ihrer Portabilität limitiert ist (Burnstein und Lederman 2003).

### **3 eKARuS – Ein webbasiertes Audience Response System**

Das *unified electronic Knowledge and Audience Response System*, kurz eKARuS wurde hauptsächlich für den Einsatz in Vorlesungen an Universitäten entwickelt. Es stellt dem Lehrenden zwei Bereiche zu Verfügung: das administrative Backend und den Präsentationsmodus. Im Backend werden Kurse und Vorlesungen für die Zuordnung der Ergebnisse angelegt, sowie der Einsatz des Systems vorbereitet. Mit dem Starten einer Vorlesungseinheit im *Präsentationsmodus* werden die zugeordneten Folien ähnlich einer Powerpoint-Präsentation gezeigt. Simultan wird ein Zugang (Live-Modus) für die Studierenden aktiviert, der über beliebige Endgeräte mit HTML5-fähigem Browser nutzbar ist.

Während der Vorlesung kann eKARuS die Interaktion zwischen Lehrenden und Lernenden durch drei Funktionalitäten unterstützen:

Die *Umfragefunktionalität* bildet das klassische ARS-Funktionsprinzip ab: Der Lehrende startet eine vorbereitete oder ad hoc erstellte Umfrage. Daraufhin wird die Abstimmungsfunktion bei den Studierenden freigeschaltet. Mit dem Beenden der Abstimmung wird die Verteilung der Antworten unmittelbar in einem Kuchendiagramm dargestellt.

*Brainstormings* können helfen, die Aufmerksamkeit der Gruppe zu binden und gelernte Inhalte zu reproduzieren oder anzuwenden (Steinert und Snell 1999). Brainstormings ähneln Umfragen, nur werden statt vorgegebener Antworten Kategorien definiert, zu denen die Studierenden Freitextantworten geben können. Die Darstellung der Ergebnisse erfolgt tabellarisch nach Kategorien, wobei die Antworten der Studierenden unmittelbar in der Ansicht des Lehrenden erscheinen.

In beiden Fällen können die Ergebnisse für die spätere Verwertung, beispielsweise für die weitere Vorlesungsplanung oder Forschung, gespeichert werden (Caldwell 2007). Ebenfalls denkbar ist eine Nutzung der Daten im Kontext von Learning Analytics, dessen Fokus u. a. auf der datenbasierten Analyse der Beziehung zwischen Lehrendem und Lernenden liegt (Long und Siemens 2011).

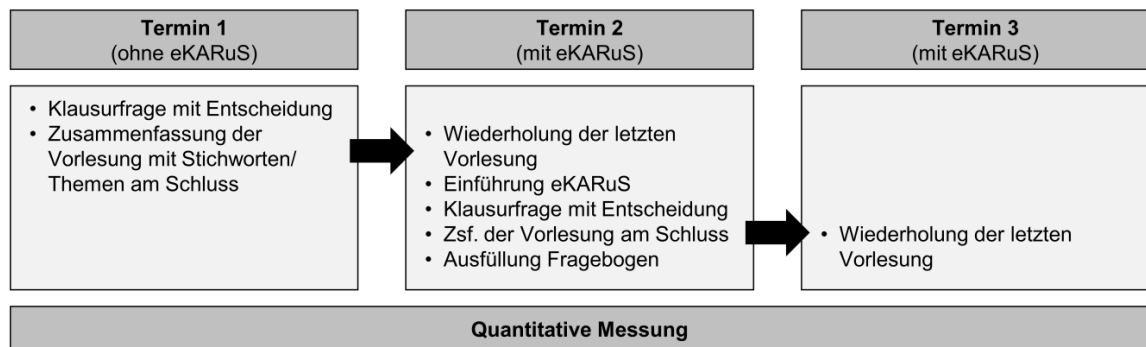
Schließlich haben die Studierenden während der kompletten Vorlesungszeit außerdem die Möglichkeit, über die *Fragen-Funktion* Fragen an den Lehrenden zu schicken. Dies kann wahlweise auch anonymisiert erfolgen.

### **4 Einsatz von eKARuS in einer universitären Vorlesung**

In Zusammenarbeit mit Dozenten der Veranstaltung „Web Applications and Internet Technology“ an der Universität Hohenheim wurde eKARuS während des laufenden Semesters in den Vorlesungsbetrieb eingeführt.

Die Vorlesung wird von Masterstudierenden besucht und auf Englisch gehalten. Aufgrund der vergleichsweise geringen Größe des Kurses von weniger als 40 Teilnehmern wurden bereits Mittel zur Erhöhung der aktiven Mitarbeit angewendet. Dazu zählen vor allem die Zusammenfassung der Inhalte der letzten Vorlesung zu Beginn sowie der aktuellen Vorlesung am Ende der Veranstaltung durch die Lernenden in Form einer offenen Diskussionsrunde.

Um eine gute Vergleichbarkeit und einen Indikator für das Behalten von Vorlesungsinhalten zu erhalten, wurden daher diese Vorgehensweisen in eKARuS abgebildet und durch weitere Vergleichspunkte ergänzt, die sich im Aufbau und Ablauf des Experiments niederschlagen. Das Experiment dauerte drei Wochen und umfasste damit drei Vorlesungstermine im gleichen zeitlichen Abstand. Eine schematische Darstellung des Aufbaus findet sich in Abbildung 1.



**Abbildung 1: Aufbau und Ablauf des Experiments (eigene Darstellung)**

An *Termin 1* wurde der Ausgangszustand durch Protokollieren des Ablaufs der Veranstaltung sowie des Interaktionsaufkommens erhoben. *Termin 2* begann noch ohne eKARuS um eine Vergleichsbasis für die wöchentliche Wiederholung der letzten Lerneinheit zu bestimmen. Abschließend wurde eKARuS mit einem einfachen, beispielhaften Anwendungsfall eingeführt. Danach nutzte der Dozent eKARuS im selben Ausmaß zur Interaktion, wie er sich traditioneller Interaktionsmittel in Termin 1 bedient hatte. Nach Ende der Vorlesungseinheit wurde ein Fragebogen für die quantitative Auswertung an die Studierenden ausgegeben. An *Termin 3* wurde zu Beginn der Vergleichszustand für die wöchentliche Wiederholung der letzten Lerneinheit mit eKARuS ermittelt. Damit endete das Klassenraumexperiment.

Die Probandengruppe bestand aus Masterstudierenden mit einem durchschnittlichen Alter von 24,78 Jahren (Standard Abweichung (SD): 1,62). Das Geschlechterverhältnis ist ausgeglichen. Alle Studierende sind in wirtschaftswissenschaftlichen Studiengängen eingeschrieben und belegen Wirtschaftsinformatik als einen von drei resp. vier Schwerpunkten, was eine durchschnittliche IT-Affinität vermuten lässt. Die Anwesenheit blieb in den ersten beiden Terminen konstant bei 32 Teilnehmern, nur in Termin 3 lag sie bei 25.

Tabelle 1 fasst die untersuchten Hypothesen zusammen. In quantitativer Hinsicht ist vor allem das *Interaktionsaufkommen* für **H2** von besonderer Wichtigkeit. Als Interaktion gilt im Rahmen dieses Experiments jede verbale oder elektronisch übermittelte Äußerung der Studierenden in Richtung des Lehrenden. Dabei wird den Interaktionen jeweils die auslösende Frage zugeordnet, um die Anzahl der *Interaktionen pro vom Lehrenden gestellte Frage* erheben zu können. Dies soll Rückschlüsse auf die Hemmschwelle zur Initiierung einer Interaktion (**H1**) zulassen.

In den Terminen 1 und 2 war jeweils eine Klausurfrage vorgesehen, um das Verständnis (**H4**) der Studierenden zu prüfen (*Anzahl korrekter Antworten zu einer Klausurfrage*).

Variable	Hypothese	Messmethode
Interaktionen pro vom Lehrenden gestellte Frage	H1	Protokollierung
Hemmschwelle für Interaktionen	H1	Fragebogen
Interaktionsaufkommen	H2	Protokollierung
Auseinandersetzen mit den Inhalten der Vorlesung	H3	Fragebogen
Anzahl korrekter Antworten zu einer Klausurfrage	H4	Protokollierung
Verständnis der Inhalte der Vorlesung	H4	Fragebogen
Behalten von Inhalten der Vorlesung	H5	Fragebogen

**Tabelle 1: Zuordnung der Messmethoden (eigene Darstellung)**

Über den Fragebogen (siehe Anhang) wurde zuerst sichergestellt, dass alle Teilnehmer das System auch verwendeten, um dann zu prüfen, mit welchem Gerät die Nutzung erfolgte. Im Anschluss wurden vier Fragen gestellt, um die subjektive Einschätzung der einzelnen Teilnehmer bezüglich der Hypothesen **H2** bis **H5** in einer Vergleichsbetrachtung mit und ohne ARS zu erfassen. Weitere Fragen erhoben potenzielle Störfaktoren wie *Ease of Use* (Davis Jr 1985), Technologie-enthusiasmus und Ablenkungspotenzial (Scheele et al. 2004). Abschließend folgten demografischen Angaben.

## 5 Evaluation

### 5.1 Auswertung der Beobachtungsprotokolle

Die Auswertung der Protokolle orientiert sich nicht an den drei Terminen, sondern an den Kategorien *ohne eKARuS* und *mit eKARuS*. Somit zählt der Wiederholungsteil vor der eKARuS-Einführung in Termin 2 zur Kategorie *ohne eKARuS*, während der Wiederholungsteil in Termin 3 zur Kategorie *mit eKARuS* zählt.

Ohne eKARuS lag das *Interaktionsaufkommen* bei 30 gezählten Interaktionen, mit eKARuS wurden insgesamt 104 Interaktionen gezählt. Dieses Ergebnis stützt die Hypothese, dass ARS das Interaktionsaufkommen erhöhen (**H2**).

Des Weiteren gab es ohne eKARuS durchschnittlich 1,07 (SD: 1,12) *Interaktionen pro vom Lehrenden gestellte Frage*, während es mit eKARuS im Mittel 6,50 (SD: 5,94) waren. Dies lässt Rückschlüsse auf die Korrektheit der Hypothese **H1** zu, dass ARS die Hemmschwelle für von Studierenden ausgehende Interaktionen in Vorlesungen herabsetzen.

Bei der *Anzahl korrekter Antworten zu einer Klausurfrage* stehen vier korrekte von fünf Abstimmungen für den Fall ohne eKARuS elf korrekten von zwölf Abstimmungen im Fall mit eKARuS gegenüber. Neben der erhöhten Teilnahme bei Verwendung von eKARuS ist damit auch ein besseres Verhältnis von korrekten zu falschen Antworten zu verzeichnen. Dies kann dafür sprechen, dass sich die Studierenden mehr mit den Inhalten der Vorlesung auseinandersetzen (**H3**) und diese besser verstehen (**H4**).

Ein Anstieg von *korrekten Beiträgen zur Wiederholung* ohne eKARuS mit zwei korrekten Beiträgen auf acht korrekte Beiträge mit eKARuS lässt Rückschlüsse darauf zu, dass sich die Verwendung von ARS positiv auf das Behalten von Inhalten auswirken könnte (**H5**).

## 5.2 Auswertung des Fragebogens

Alle Studierende (n=32) verwendeten eKARuS während des Experiments. Die Hälfte der Teilnehmer nutzte das System per Smartphone, der Rest überwiegend per Laptop (n=12). Bis auf drei Ausnahmen wurden die Fragebögen vollständig ausgefüllt. Der Datensatz enthält darüber hinaus keine außergewöhnlichen Ausreißer.

Die Ergebnisse der Fragen zu den Faktoren des Lernerfolgs werden auf Basis einer 7-Punkte Likert-Skala ausgewertet, deren mittlerer Wert als 0 bzw. kein Effekt zu interpretieren ist. Entsprechend ist das erste Feld (+3) als „positiver Effekt“ und das letzte Feld als „negativer Effekt“ (-3) zu werten.

	N	Minimum	Maximum	Mittelwert	Standardabweichung
Hemmschwelle für Interaktionen	31	-3	3	,87	1,89
Auseinandersetzen mit den Inhalten der Vorlesung	32	-3	3	1,62	1,47
Verständnis der Inhalte der Vorlesung	32	-3	3	,71	1,57
Behalten von Inhalten der Vorlesung	32	-2	3	,96	1,44
Gültige Anzahl (listenweise)	31				

**Tabelle 2: Deskriptive Statistiken zum Lernerfolg**

Die Auswertung zeigt ein ambivalentes Bild: Bei der Variablen *Auseinandersetzen mit Vorlesungsinhalten* (vgl. mit *traditioneller Vorlesung*) wird mit einem Mittelwert von 1,63 (SD: 1,48) ein positiver Effekt des Artefakts wahrgenommen. Dagegen sprechen sich zwei Teilnehmer für einen negativen Effekt aus. Im Gesamtbild stützt dieses Ergebnis die Hypothese **H3**.

Beim *Verständnis von Vorlesungsinhalten* zeigt sich die Ambivalenz besonders deutlich: Hier wird der Effekt von eKARuS von den Lernenden stark individuell unterschiedlich wahrgenommen (M: 0,72 SD: 1,57). Ähnlich verhält es sich auch mit dem *Merken von Vorlesungsinhalten* (M: 0,97/SD: 1,45) und der *subjektiven Hemmschwelle für Interaktionen* (M: 0,87/SD: 1,45) im Vergleich mit traditionellen Vorlesungen. Für die Hypothesen H2, H4 und **H5** bedeutet dies zwar eine Tendenz ins Positive, die Heterogenität der Resultate stellt deren Aussagekraft aber grundlegend in Frage.

Um die Unabhängigkeit dieser Ergebnisse von beeinflussenden Faktoren abzusichern, wurden im weiteren Verlauf des Fragebogens Fragen gestellt, die auf diese Effekte abzielen. Da es sich dabei nicht wie zuvor um eine Vergleichsbetrachtung handelt, wird hier ein Intervall von 1 bis 7 angesetzt, wobei 1 mit voller Zustimmung und 7 mit absoluter Ablehnung gleichzusetzen ist.

	N	Minimum	Maximum	Mittelwert	Standardabweichung
Ease of Use	32	1	2	1,25	,43
Technologieenthusiasmus	32	3	7	4,81	1,35
Ablenkung durch eKARuS	32	1	7	4,93	1,86
Gültige Anzahl (listenweise)	32				

**Tabelle 3: Deskriptive Statistiken zu beeinflussenden Faktoren**

Die Nutzung des Systems wird mit einem Mittelwert von 1,25 und einer geringen Standardabweichung von 0,44 durchgehend als einfach eingestuft. Der Einfluss von Technologieenthusiasmus wird von den Teilnehmern als eher gering bewertet (M: 4,81/SD: 1,35). Des Weiteren schätzen die Probanden die Ablenkung durch eKARuS mit einem Mittelwert von 4,94 (SD: 1,87) ebenfalls als weniger relevant ein.

## 6 Diskussion

### 6.1 Limitationen

Die Stichprobe der Lernenden, die an dem Experiment teilgenommen hat, ist neben ihrer geringen Größe auch in ihrer Zusammensetzung speziell. Da über die drei Termine nur diese eine Gruppe beobachtet wurde, fehlt es an Vergleichswerten, um die Ergebnisse auf einen größeren Teil der Lernenden beziehen zu können. Auch über die drei Termine ist die konstante Anwesenheit aller Teilnehmer nicht garantiert. Es handelt sich also um die Betrachtung einer weder vollständig homogenen noch ganzheitlich heterogenen Versuchsgruppe. Zwar spiegelt dies den realen Zustand im Vorlesungskontext eher wieder als die in Laborexperimenten geforderte genaue Zuordnung der Probanden, etwaige Drittvariableneffekte können die Ergebnisse allerdings stärker verfälschen.

Durch die begrenzten Zeitrahmen war nur ein kurzer Einblick in den Ist-Zustand der traditionellen Vorlesung im Rahmen eines Vorlesungstermins möglich. Die darin gewonnenen quantitativen Daten müssen daher auch als Ausschnitt bzw. als Einzelfall betrachtet werden, da exogene Effekte auf die Motivation, Teilnahmebereitschaft und tatsächliche Interaktionen bei fehlendem Vergleich nicht ausgeschlossen werden können. Gleiches gilt für die Termine mit eKARuS. Vor allem in Zusammenhang mit der Einführung des Systems und der Neuartigkeit des Vorlesungsablaufs kann Technologieenthusiasmus die Ergebnisse verzerren. Es sei an dieser Stelle allerdings auf ähnliche Experimente verwiesen, die zeigen, dass dieser keinen allzu lang anhaltenden Effekt besitzt und weitestgehend vernachlässigt werden kann (Draper und Brown 2004; Scheele et al. 2004). Auf der anderen Seite nimmt die Effektivität von ARS nach Draper und Brown (2004) über die Zeit zu und es relativiert sich die Wahrscheinlichkeit, weniger Stoff in der Vorlesungseinheit vermitteln zu können. Grund hierfür ist, dass die Anwendung wiederholt wird und dadurch Übungseffekte erzielt werden (Dufresne et al. 1996). Diese Effekte konnten während des kurz bemessenen Einsatzes ebenfalls nicht wirken.

Ein weiterer Punkt, dessen Einfluss nicht auszuschließen ist, ist die Heterogenität der Vorlesungsinhalte über die zwei Vergleichstermine. Zwar wurde darauf geachtet, inhaltlich ähnliche und stark zusammenhängende Themen zu behandeln, Unterschiede in der Verständnis- und Merkschwierigkeit können jedoch nicht mit Sicherheit ausgeschlossen werden.

Schlussendlich gab es auch im operativen Ablauf des Experiments Vorkommnisse, die die Validität der Ergebnisse beeinflussen können. So traten in Verbindung mit eKARuS an Termin 2 und Termin 3 unvorhersehbare technische Störungen in der Netzwerkinfrastruktur auf.

### 6.2 Folgerungen und Einordnung der Ergebnisse

Zielsetzung des Experiments und der gewählten Versuchsanordnung war es, am Beispiel von eKARuS den Einfluss von ARS auf den Lernerfolg von Studierenden im mobilen Informationszeitalter in einem Realwelt-Szenario zu bestimmen. Die aufgestellten Hypothesen konnten nur teilweise validiert werden.

Die quantitativen Analysen bestätigen, dass ARS das Interaktionsaufkommen erhöhen (**H2**). Ausgehend von der Theorie des aktiven Lernens (Bonwell und Eison 1991) bzw. der Strömung der interaktiven Vorlesung müsste mit diesem eine stärkere Auseinandersetzung mit den Vorlesungsinhalten seitens der Studierenden (**H3**) einhergehen. Dies sollte, nach bisherigen Erkenntnissen (s. Kapitel 2), zu einem besseren Verständnis (**H4**) und Behalten (**H5**) der Inhalte führen.

Die erste Folgerung in Bezug auf das Auseinandersetzen mit den Vorlesungsinhalten zeigt bereits in der Protokollauswertung eine Tendenz ins Positive, welche durch die Fragebögen belegt werden.

In Bezug auf **H4** und **H5** zeigen die Protokolldaten zwar ebenfalls positive Ergebnisse, die Fragebogenergebnisse ergeben aber kein ausreichend eindeutiges Bild. Dies widerspricht den Ergebnissen früherer Experimente, welche größtenteils positive, zumindest aber in ihrer Tendenz eindeutige Auswirkungen beschreiben (Dufresne et al. 1996).

Ähnlich verhält es sich mit den Beobachtungen bezüglich der Hemmschwelle zur Interaktionsinitiierung bei der Verwendung von ARS (**H1**). Zwar sprechen die Ergebnisse der quantitativen Analyse für eine positive Bewertung dieser Hypothese, von den Lernenden konnte dieses Phänomen aber nicht eindeutig bestätigt werden.

In diesem Zusammenhang sind weitere Untersuchungen in Hinblick auf die Gründe der Ambivalenz der hier erhaltenen Ergebnisse notwendig, um verlässliche Aussagen treffen zu können. Als Ausgangspunkt können hierfür die Limitationen (s. Kapitel 6.1) dienen. Vor allem die begrenzte Zeit in Verbindung mit der Herausforderung der Integration von eKARuS in einer pädagogisch zielführenden Art und Weise kann hier als prominente Einflusskraft wirken. In der kurzen Betrachtungsperiode konnte der Einsatz des ARS noch keinen ausgereiften, durch Erfahrung optimierten Routinestand erreichen. Dadurch wurde die Verwendung möglicherweise von bestimmten Studierenden als nicht effektiv eingeschätzt. Dieses kritische Hinterfragen der pädagogischen Legitimation wurde bei den Untersuchungen von Draper und Brown (2004) ebenfalls beobachtet.

Zu den potenziellen Störfaktoren ist zu sagen, dass die Nutzungsschwierigkeit einen wenig relevanten Einflussfaktor darstellt. Zwar scheint dies auf den ersten Blick auch für das Ablenkungspotenzial bei ARS zu gelten, hier sollte jedoch bedacht werden, dass diese Bewertung seitens der Studenten auch auf der mit der Nutzung des Systems einhergehenden Legitimation zur Nutzung des Smartphones basieren könnte.

## 7 Fazit

Der Wandel zu einer ständig vernetzten Gesellschaft, die sich täglich einer Flut von Informationen ausgesetzt sieht, wirkt sich auf eine Vielzahl verschiedener Bereiche aus, auch, oder vielleicht im besonderen Maße, auf die Lehre. Das Hauptziel, den Lernerfolg zu steigern, bleibt dasselbe.

Inwiefern zeitgemäße ARS dies erreichen wurde auf quantitativer und qualitativer Ebene untersucht. Zwar wurden erhöhtes Interaktionsaufkommen und verstärkte Auseinandersetzung mit den Vorlesungsinhalten beobachtet, die Resultate bezüglich des Verständnisses und des Behaltens der Inhalte zeigen jedoch ein ambivalentes Bild. Die möglichen Gründe dafür sind vielfältig, aber schwer einzugrenzen.

Auffällig ist jedoch der Unterschied zu Ergebnissen früherer Experimente ähnlicher Natur: Sollte die Abweichung nicht in den Limitationen des Versuchsaufbaus begründet sein, so führt dies zu der Vermutung, dass sich der Wandel zur Informationsgesellschaft und dem damit einhergehenden veränderten Lernverhalten der Menschen auf die Eignung von Werkzeugen wie ARS auswirken, was weiteren Forschungsbedarf induziert.

## 8 Anhang

Questionnaire						Date	
In this lecture you have been introduced to the prototype system eKARuS. Please read the following statements carefully and mark with a cross on the right side how much you would agree to them. The information provided will be gathered anonymously.							
I have used eKARuS in today's lecture.				<input type="checkbox"/> yes		<input type="checkbox"/> no	
[if 'yes':] I have used eKARuS via the following device:				<input type="checkbox"/> Smartphone <input type="checkbox"/> Tablet <input type="checkbox"/> Laptop <input type="checkbox"/> Other device: _____			
	strongly agree	agree	agree somewhat	no difference	disagree somewhat	disagree	strongly disagree
When using eKARuS, I engage myself in the content of the lecture more than in lectures without eKARuS.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
When using eKARuS, I understand the content of the lecture better than in lectures without eKARuS.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
When using eKARuS, I can memorize the content of the lecture better than in lectures without eKARuS.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
When using eKARuS, the inhibition level to participate actively in the lecture is lower than in lectures without eKARuS.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	strongly agree	agree	agree somewhat	neutral	disagree somewhat	disagree	strongly disagree
I find eKARuS easy to use.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The use of eKARuS would be boring for me very quickly.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The use of eKARuS distracts me from the actual content of the lecture.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Abbildung 2: Originaler Fragebogen (ohne demographische Daten)

## 9 Literatur

- Anzai Y, Simon HA (1979) The theory of learning by doing. *Psychological Review* 86:124–140.
- Bishop JL, Verleger MA (2013) The flipped classroom; A survey of the research. In: American Society for Engineering Education (Hrsg) ASEE National Conference Proceedings, Atlanta.
- Bonwell CC, Eison JA (1991) Active learning; Creating excitement in the classroom. George Washington University, Washington, D.C.
- Boyle JT, Nicol DJ (2003) Using classroom communication systems to support interaction and discussion in large class settings. *Association for Learning Technology Journal* 11:43–57.
- Burnstein RA, Lederman LM (2003) Comparison of Different Commercial Wireless Keypad Systems. *The Physics Teacher* 41:272–275.
- Butler JA (1992) Use of Teaching Methods within the Lecture Format. *Medical Teacher* 14:11–25.
- Cain J, Robinson E (2008) A Primer on Audience Response Systems: Current Applications and Future Considerations. *American Journal of Pharmaceutical Education* 72:1–6.
- Caldwell JE (2007) Clickers in the Large Classroom; Current Research and Best-Practice Tips. *Cell Biology Education* 6:9–20.
- Cochrane TD (2010) Exploring Mobile Learning Success Factors. *Research in Learning Technology* 18:133–148.
- Crouch CH, Mazur E (2001) Peer Instruction; Ten years of experience and results. *American Journal of Physics* 69:970–977.
- Davis Jr FD (1985) A technology acceptance model for empirically testing new end-user information systems; Theory and Results. Dissertation, Cambridge.



- Doise W, Mugny G, Perret-Clermont A (1975) Social interaction and the development of cognitive operations. *European Journal of Social Psychology* 5:367–383.
- Draper SW, Brown MI (2004) Increasing interactivity in lectures using an electronic voting system. *Journal of Computer Assisted Learning* 20:81–94.
- Draper SW, Cargill J, Cutts Q (2002) Electronically enhanced classroom interaction. *Australian Journal of Educational Technology* 18:13–23.
- Dufresne RJ, Gerace WJ, Leonard WJ, Mestre JP, Wenk L (1996) Classtalk: A classroom communication system for active learning. *Journal of Computing in Higher Education* 7:3–47.
- Efstathiou N, Bailey C (2012) Promoting active learning using Audience Response System in large bioscience classes. *Nurse Education Today* 32:91–95.
- Fies C, Marshall J (2006) Classroom Response Systems; A Review of the Literature. *Journal of Science Education and Technology* 15:101–109.
- Froehlich HP (1963) What about classroom Communicators? *Audiovisual communication review* 11:19–26.
- Garrison DR (2011) *E-learning in the 21st century; A framework for research and practice*. Routledge, New York.
- Gartner (2015) Number of smartphones sold to end users worldwide from 2007 to 2014 (in million units). <http://www.statista.com/statistics/263437/global-smartphone-sales-to-end-users-since-2007/>. Zugegriffen: 02. Juli 2015.
- Graham CR, Tripp TR, Seawright L, Joeckel GL (2007) Empowering or compelling reluctant participants using audience response systems. *Active Learning in Higher Education* 8:233–258.
- Helmerich J, Scherer J (2007) Interaktion zwischen Lehrenden und Lernenden in Medien unterstützten Veranstaltungen. In: Breitner MH, Bruns B, Lehner F (Hrsg) *Neue Trends im E-Learning. Aspekte der Betriebswirtschaftslehre und Informatik*. Physica-Verlag Heidelberg, Heidelberg, S 197–210.
- Kay RH, LeSage A (2009a) A strategic assessment of audience response systems used in higher education. *Australasian Journal of Educational Technology* 25:235–249.
- Kay RH, LeSage A (2009b) Examining the benefits and challenges of using audience response systems; A review of the literature. *Computers & Education* 53:819–827.
- Latessa R, Mouw D (2005) Use of an audience response system to augment interactive learning. *Family Medicine* 37:12–14.
- Long P, Siemens G (2011) Penetrating the fog; Analytics in Learning and Education. *EDUCAUSE Review*, 46:31–40.
- Naismith L, Lonsdale P, Vavoula GN, Sharples M (2004) Literature review in mobile technologies and learning; A report for NESTA Futurelab. NESTA Futurelab, Bristol.
- Nichols M (2003) A theory for eLearning. *Journal of Educational Technology & Society* 6:1–10.
- Schackow E, Chavez M, Loya L, Friedmann M (2004) Audience Response System; Effect on Learning in Family Medicine Residents. *Family Medicine* 36:496–504.
- Scheele N, Wessels A, Effelsberg W (2004) Die Interaktive Vorlesung in der Praxis. In: Engels G, Seehusen S (Hrsg) *DeLFI 2004. Die 2. e-Learning Fachtagung Informatik*. Gesellschaft für Informatik, Bonn, S 283–294.
- Schmidt B (2011) Teaching engineering dynamics by use of peer instruction supported by an audience response system, *European Journal of Engineering Education* 36:413–423.
- Springer L, Stanne ME, Donovan SS (1999) Effects of Small-Group Learning on Undergraduates in Science, Mathematics, Engineering, and Technology: A Meta-Analysis. *Review of Educational Research* 69:21–51.
- Steinert Y, Snell LS (1999) Interactive lecturing; Strategies for increasing participation in large group presentations. *Medical Teacher* 21:37–42.
- Tregonning AM, Doherty DA, Hornbuckle J, Dickinson JE (2012) The audience response system and knowledge gain; A prospective study. *Medical Teacher* 34:269–274.
- Vygotskiï LS, Cole M (1978) *Mind in society; The development of higher psychological processes*. Harvard University Press, Cambridge.
- Wenz HJ, Zupanic M, Klosa K, Schneider B, Karsten G (2014) Using an audience response system to improve learning success in practical skills training courses in dental studies; a randomised, controlled cross-over study. *European Journal of Dental Education* 18:147–153.



# **Fraud Detection im Gesundheitswesen mithilfe von Data-Mining-Clusterverfahren**

**Raoul Könsgen<sup>1</sup> und Steffen Stock<sup>1</sup>**

<sup>1</sup> Europäische Fachhochschule, Fachbereich Wirtschaftsinformatik,  
raoul.koensgen@eufh-mail.de, s.stock@eufh.de

## **Abstract**

Die Abrechnungsbetrugskosten im Gesundheitswesen beliefen sich im Jahr 2014 auf circa 21 Milliarden Euro (Statistisches Bundesamt 2015). Die Mehrausgaben führen dabei zu steigenden Beiträgen für Krankenversicherungen. Die Konvergenz von neuen Kommunikationstechnologien und Innovationen aus dem Bereich der Telemedizin stimuliert den Trend zu Big Data. Insbesondere das Datenvolumen im Gesundheitswesen ist hiervon betroffen und es wird zunehmend schwieriger, wertvolle Informationen zur Aufdeckung von Abrechnungsbetrug aus der Datenflut zu extrahieren. Ein Fraud-Detection-System zur Aufdeckung von Abrechnungsbetrug auf Basis von Data-Mining-Clusterverfahren könnte dazu beitragen, die Kosten im Gesundheitswesen zu senken. Ziel ist es zu prüfen, wie Data-Mining-Clusterverfahren zur Aufdeckung von Abrechnungsbetrug im Gesundheitswesen genutzt werden können. Hierzu werden zwei Vorgehen mit ausgewählten Algorithmen beschrieben, um Abrechnungsbetrug aufzudecken. Anschließend wird, unter Berücksichtigung der Zielsetzung, das Vorgehen einer Datenanalyse im Rahmen des Knowledge-Discovery-in-Databases-Prozesses aufgezeigt.

## **1 Einleitung**

Die Anwendungsmöglichkeiten von Fraud im Gesundheitswesen sind vielfältig und umfassen neben einer Abrechnungsfälschung ebenfalls Rezeptfälschung und Chipkartenmissbrauch. Nach Ansicht des Bundeskriminalamtes handelt es sich bei Abrechnungsfälschung um eine besonders sozialschädliche Form der Wirtschaftskriminalität, da die Integrität des Gesundheitswesens negativ beeinflusst wird (Bundeskriminalamt 2004). Die Ärzte gehören mit einem Anteil von 14,6 Prozent zur Leistungserbringergruppe, die am häufigsten unter Tatverdacht steht (Pricewaterhouse Cooper 2012). Jede Krankenkasse ist gesetzlich verpflichtet, eine Stelle zur Bekämpfung von Fehlverhalten im Gesundheitswesen einzurichten (§197a SGB V). Darüber hinaus kann ein Unternehmen mithilfe einer erfolgreichen Schadensbekämpfung einen Wettbewerbsvorteil erlangen.

Für eine korrekte Abrechnung der deutschen Hausärzte, im Rahmen der gesetzlichen Krankenversicherung ist die Kassenärztliche Vereinigung (KV) und der Deutsche Hausärzterverband e.V. verantwortlich. Jedes dieser Unternehmen verfügt über eigene Rechenzentren. Die Abrechnungs-

daten werden von den Arztinformationssystemen an die Rechenzentren übermittelt, geprüft, transformiert und an die jeweilige Krankenkasse weitergeleitet.

In der Literatur werden keine konkreten Vorgehensweisen zur Aufdeckung von Fraud mithilfe von Data-Mining-Clusterverfahren aufgeführt (Hansen et al. 2015). Aus diesem Grund soll in die Frage beantwortet werden, welche möglichen Vorgehensweisen zur Aufdeckung von Fraud mithilfe von Data-Mining-Clusterverfahren eingesetzt werden. Unter Berücksichtigung der Zielsetzung, werden anhand eines Fallbeispiels alle notwendigen Schritte einer Datenanalyse erläutert.

Zu Beginn werden in Kapitel 2 die Grundlagen von Fraud Detection und Data-Mining aufgezeigt. Anschließend werden in Kapitel 3 zwei Vorgehen zur Aufdeckung von Fraud mithilfe Data-Mining-Clusterverfahren aufgezeigt. Auf Basis dieser Ergebnisse wird in Kapitel 4 eine Datenanalyse zur Aufdeckung von Abrechnungsbetrug beschrieben. Ein Fazit bildet den Abschluss in Kapitel 5.

## 2 Grundlagen Fraud Detection und Data-Mining

Im Folgenden werden die Begriffe Fraud, Fraud Detection und Data-Mining definiert. Anschließend werden die beiden letztgenannten Begriffe in einen Zusammenhang gebracht.

In der Literatur gibt es keine einheitliche Begriffsdefinition für Fraud. Es wird häufig auf zwei Definitionen verwiesen (Nimwegen 2009). Es handelt sich um die Ausführungen des Instituts der Wirtschaftsprüfer Deutschlands e. V. (IWD 2010) und der Association of Certified Fraud Examiners (Association of Certified Fraud Examiners 2015). Das IWD definiert Fraud als beabsichtigte Verstöße gegen Rechnungslegungsgrundsätze, die zu falschen Angaben im Abschlussbericht führen. Zusätzlich wird der Begriff Fraud in die Kategorien Täuschung, Vermögensschädigung und Gesetzesverstöße unterteilt. Unter dem Begriff der Täuschung ist die falsche Anwendung in der Rechnungslegungsvorschriften und das unbefugte Ändern in der Buchführung gemeint. Vermögensschädigungen bezeichnen Handlungen, die widerrechtlich auf das Aneignen oder Vermindern des Gesellschaftsvermögens ausgerichtet sind. Ein Gesetzesverstoß liegt vor, wenn ein Unternehmen z. B. mögliche Schadensersatzverpflichtungen oder Geldstrafen nicht in der Rechnungslegung berücksichtigt (Nimwegen 2009). Die ACFE unterliegt keiner strengen juristischen, sondern folgt einer wirtschaftlichen Betrachtungsweise. Es verweist auf folgende Definition. „A knowing misrepresentation of the truth or concealment of a material fact to induce another act to his or her detriment“ (Association of Certified Fraud Examiners 2015). Das Fraud-Triangle ist ein Instrument zur Risikobewertung und beschreibt die Neigungen zu Fraud unter Berücksichtigung der Faktoren: Anreiz / Druck, Gelegenheit und innerliche Rechtfertigung, welche im Folgenden beschrieben werden (Cressey 1973):

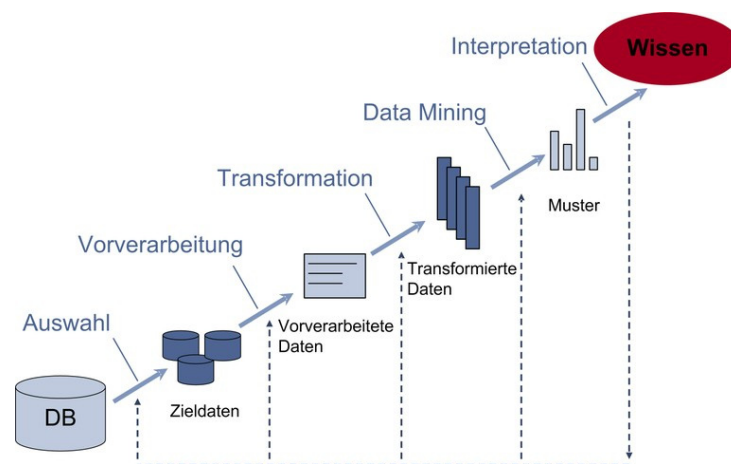
- Der Faktor Anreiz / Druck gibt den Grund an, warum Täter wirtschaftskriminell handeln. Ein häufiger Grund sind finanzielle Verpflichtungen, denen nicht nachgekommen werden kann. Darüber hinaus hat sich gezeigt, dass der Täter diesen Faktor selten mit keiner dritten Person teilt, da nach dessen Meinung dies mit einer Schwächung des gesellschaftlichen Bildes verbunden sein würde.
- Eine Gelegenheit setzt voraus, dass der Täter den gewährten Handlungsspielraum nutzen kann und zusätzlich über ausreichende Fähigkeiten verfügt. Schwächen im Kontrollsystem führen dabei zu Gelegenheiten. Es wird häufig ein systematisches Vorgehen verfolgt, bei dem der Täter das Kontrollsystem testet, indem er z. B. einen unwesentlichen Betrag von einem Konto abhebt.

- Eine innerliche Rechtfertigung resultiert aus der Tatsache, dass der Täter sein Vorgehen als nicht kriminell ansieht oder sich nicht für sein Vorgehen verantwortlich fühlt. Die Gefahr von „moral hazard“ kann durch Anreiz- und Kontrollsysteme reduziert werden.

Fraud-Detection umfasst adäquate und effektive Maßnahmen zur Aufdeckung von fraudulenten Handlungen. Es wird ein Ex-Post-Ansatz verfolgt, der auf eine vergangenheitsorientierte Betrachtung ausgelegt ist (Baesens et al. 2015). Fraud Detection setzt an dem Faktor Gelegenheit des Fraud-Triangles an und versucht diesen zu minimieren.

Der Begriff Data-Mining hat einen populärwissenschaftlichen Ursprung und ist aus der Metapher Bergbau abgeleitet, dessen Gegenstand der Abbau von Kohle in Erdmassen darstellt. Es handelt sich um ein interdisziplinäres Forschungsgebiet mit den angrenzenden Bereichen der Informatik, Mathematik und Statistik (Petersohn 2005). Data-Mining umfasst die Extraktion und Entdeckung von implizitem, bisher nicht bekanntem, nicht-trivialem und potentiell nützlichem Wissen aus Daten (Frawley et al. 1992).

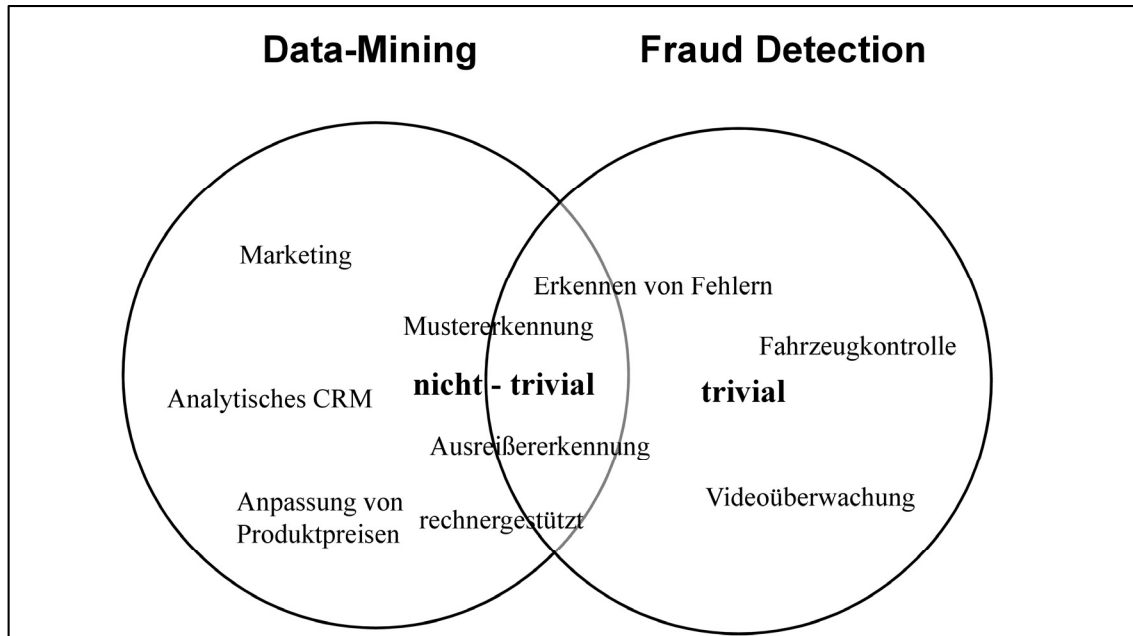
Es existiert kein allgemeingültiger Prozess zur Datenanalyse (Freiknecht 2014 und Sharafi 2013). Gemeinsam haben alle Modelle, dass es sich um einen Prozess zur Identifizierung von nichttrivialen, unbekanntem und potentiell nützlichen Informationen aus Daten handelt. Der Prozessablauf ist iterativ, interaktiv und umfasst mehrere Teilprozesse. Das Forschungsgebiet Knowledge Discovery in Databases (KDD) hat das Ziel, in komplexen Datenbeständen implizit vorhandenes Wissen aufzudecken und explizit zu machen (Chen et al. 1996). KDD lässt sich als Prozess mit den Phasen Auswahl, Vorverarbeitung, Transformation, Data-Mining und Interpretation gemäß Bild 1 darstellen.



**Bild 1: Knowledge Discovery in Databases Prozess (Fayyad et al. 1996)**

Eine Gegenüberstellung von Data-Mining und Fraud Detection ist in Bild 2 dargestellt. Es wird deutlich, dass Fraud Detection aus nicht-trivialen und trivialen Maßnahmen bestehen kann. Beispielsweise ist eine Fahrzeugkontrolle eine Maßnahme von Fraud Detection von der Data-Mining nicht betroffen ist. Data-Mining umfasst nur den nicht-trivialen Themenbereich, in dem z. B. Analytische Customer-Relationship-Management-Systeme eingesetzt werden. Zur Aufdeckung von Fraud kann Data-Mining genutzt werden, um aus bekannten fraudulenten Handlungen ein Muster abzuleiten. Diese Muster können verwendet werden, um weitere Täter im Datenbestand aufzudecken. Unter der Annahme, dass Täter von den durchschnittlichen Eigenschaften abweichen, können Verfahren zur Ausreißererkennung genutzt werden. Data-Mining lässt sich auf

verschiedenen Datenbeständen anwenden und ist unabhängig von der Branche. Neben der Aufdeckung von Abrechnungsbetrug, wird Data-Mining beispielsweise bereits erfolgreich von Banken zur Erkennung von Kreditkartenmissbrauch eingesetzt.



**Bild 2: Gegenüberstellung Fraud Detection und Data-Mining**

### 3 Data-Mining-Clusterverfahren für Fraud Detection

Die Clusterverfahren gehören neben den Künstlich Neuronalen Netzen, Entscheidungsbäumen und der Assoziationsanalyse zu den Verfahren des Data-Minings (Chamoni et al. 2010). Grundsätzlich können alle Verfahren für Fraud Detection eingesetzt werden, da die Suche nach Auffälligkeiten oder Gemeinsamkeit unabhängig vom Verfahren ist. Zusätzlich ist es möglich, dass verschiedene Perspektiven der Modellierung eingenommen werden. Dies bedeutet, dass fraudulente Handlungen auf Basis einer Gesamtheit von Transaktionen oder einzelnen Transaktionen aufgedeckt werden. Im Rahmen der Arbeit wird erste Perspektive eingenommen. Im Folgenden werden die Einsatzmöglichkeiten von Data-Mining-Clusterverfahren für Fraud Detection diskutiert. Hierzu werden zwei verschiedene Vorgehen erläutert, um Ärzte, die Abrechnungsbetrug begehen, im Folgenden Fraud-Ärzte genannt, zu identifizieren.

#### 3.1 Vorgehen 1

Dieses Vorgehen basiert auf der Annahme, dass bekannte Fraud-Ärzte ein gemeinsames Verhaltensmuster aufweisen und somit auf weitere unbekannte Fraud-Ärzte geschlossen werden kann. Hierzu eignet sich eine Clusteranalyse, bei der die Objekte die Ärzte sind. Ziel ist es, dass der entsprechende Clusteralgorithmus trainiert wird, dass sich bekannte Fraud-Ärzte sich im selben Cluster befinden, im Folgenden Fraud-Cluster genannt. Die Anzahl der Cluster soll dabei möglich hoch sein, um Verhaltensmuster von Fraud-Ärzten möglichst genau zu spezifizieren. Ärzte, bei denen bisher nicht bekannt ist, ob es sich um Fraud-Ärzte handelt und sich im Fraud-Cluster befinden, weisen das selbe Verhaltensmuster, wie bekannte Fraud-Ärzte auf und sind daher mit einer höheren Wahrscheinlichkeit Fraud-Ärzte. Insbesondere partitionierende Clusterverfahren

eignen sich hierzu, da die Anzahl der Cluster zu Trainingszwecken selbst festgelegt werden kann und jedem Objekt mindestens ein Cluster zugewiesen wird. Für dieses Vorgehen kann der k-Means-Algorithmus verwendet werden, da dieser eine schnelle Laufzeit hat und die genannten Anforderungen erfüllt.

Das Ziel von k-Means ist es, den Datensatz in  $k$  Cluster aufzuteilen, sodass die Summe der quadrierten Abweichungen vom Mittelpunkt des Clusters minimal ist. Es handelt sich um eine Vektorquantisierungsmethode, welche die Zuordnung von Datenpunkten zu den nächstgelegenen Mittelpunkten eines Clusters vornimmt (MacQueen 1967). Diese basiert implizit auf dem Distanzmaß der Euklidischen Distanz, weil die Summe der quadrierten Abweichungen vom Mittelpunkt gleich der Summe der paarweise quadrierten Euklidischen Distanz, geteilt durch die Anzahl der Datenpunkte ist.

Die Euklidischen Distanz zwischen den Punkten  $x$  und  $y$  ist definiert als:

$$d(x, y) = \sqrt{\sum_{j=1}^n (x_j - y_j)^2} . \quad (1)$$

Für die folgende Zielfunktion von k-Means seien  $x_j$  die Datenpunkte und  $\mu_j$  die Mittelpunkte der jeweiligen Cluster  $S_i$ .

$$J = \sum_{i=1}^k \sum_{x_j \in S_i} \|x_j - \mu_j\|^2 . \quad (2)$$

Der Lloyd-Algorithmus ist der am häufigsten verwendete Algorithmus zur Minimierung von (2). Im Folgenden werden die Schritte des Lloyd-Algorithmus beschrieben. Im ersten Schritt werden  $k$  zufällige Werte aus dem Datensatz  $x_j$  gewählt. Die Anzahl der Iterationen werden durch  $t$  festgelegt. Im zweiten Schritt werden alle umliegenden Datenpunkte dem Cluster mit dem nächstgelegenen Mittelpunkt zugeordnet. Die Mittelpunktberechnung bestimmt die Position des Mittelpunkt und ergibt sich aus:

$$m_i^{(t+1)} = \frac{1}{|S_i^{(t)}|} \sum_{x_j \in S_i^{(t)}} x_j . \quad (3)$$

Im dritten Schritt werden die Mittelpunkte der neuen Cluster berechnet und es erfolgt eine erneute Zuordnung der Datenpunkte zu dem Clustern mit dem jeweils geringsten Abstands zum Mittelpunkt. Die Schritte eins bis drei werden wiederholt, bis die Anzahl der vorgegebenen Iterationen erreicht ist.

### 3.2 Vorgehen 2

Dieses Vorgehen geht davon aus, dass Fraud-Ärzte sich anders verhalten als korrekt abrechnende Ärzte und über Ausreißerverfahren identifiziert werden können. Die Verhaltensmuster von Fraud-Ärzten weichen deutlich vom durchschnittlichen Verhaltensmuster der Gesamtheit ab. Es eignen sich Ausreißerverfahren, um auf bisher unbekannte Fraud-Arzte zu schließen. Hierzu wird entsprechende Algorithmus zur Ausreißerkennung mit dem Ziel trainiert, dass bekannte Fraud-Ärzte einen relativ hohen Ausreißerwert aufweisen. Insbesondere die Ansätze zur Erkennung von Ausreißern auf Basis von dichte-basierte Clusterverfahren eignen sich für dieses Vorgehen, da die

Dichte eines Objektes über den durchschnittlichen Abstand zu den nächsten  $n$  Objekten definiert ist.

Ansätze zur Erkennung von Ausreißern auf Basis zur Gesamtheit aller Objekte haben nicht die Möglichkeit, Ausreißer gegenüber einer Teilmenge von Objekten zu identifizieren. Als Lösungsansatz kann der Local Outlier Factor (LOF) herangezogen werden (Breunig et al. 2000). Der LOF basiert auf dem Density-Based Spatial Clustering of Applications with Noise (DBSCAN) und ist ein dichtebasierter Algorithmus zur Erkennung von Ausreißern. Hierzu wird jedem Objekt ein LOF zugewiesen, der bestimmt, nach welchem Grad ihm ein Outliner-Status zugeordnet werden kann. Hierbei werden Ausreißer nicht nur qualifiziert, sondern auch quantifiziert. Je geringer der LOF, desto geringer ist die Wahrscheinlichkeit für einen Ausreißer. Im Folgenden werden die Bestandteile des LOF beschrieben:

Bei einer gegebenen Datenmenge  $X$  sei für jede positive Zahl  $k$  die  $k$ -Distanz ( $x$ ) eines Objektes  $x$  für jede positive Zahl definiert, als Entfernung  $d(x, x')$  zwischen  $x$  und  $x'$  aus  $X$ . Dabei gilt für mindestens  $k$  Objekte  $x'' \in X \setminus \{x\}$ , dass  $d(x, x'') \leq d(x, x')$  und für maximal  $k-1$  Objekte  $x''' \in X \setminus \{x\}$  gilt  $d(x, x''') < d(x, x')$ .

Die  $k$ -Distanz Nachbarschaft umfasst bei gegebener  $k$ -Distanz ( $x$ ) die Menge von Objekten  $x'$ , deren Distanz von  $x$  nicht größer als die  $k$ -Distanz ( $x$ ) ist. Diese ist definiert als:

$$N_k(x) = \{x' \in X \setminus \{x\} \mid d(x, x') \leq k_d(x)\} . \quad (4)$$

Die Erreichbarkeitsdistanz  $rd$  eines Objektes  $x$  zu dem Objekt  $x'$  mit  $k \in \mathbb{N}$ , sei definiert als:

$$rd_k(x, x') = \max\{k_d(x'), d(x, x')\} . \quad (5)$$

Die Erreichbarkeitsdistanz von  $x$  zu  $x'$  ist somit der wahre Abstand, außer wenn die tatsächliche Distanz kleiner als  $k$ -Distanz  $x'$  ist, dann wird die  $rd_k(x, x')$  durch die  $k$ -Distanz von  $x'$  ersetzt. Dieses Vorgehen ist dadurch begründet, dass die statistische Fluktuationen für die Objekte  $x$ , welche eine kurze Distanz zu  $x'$  haben, deutlich reduziert wird.

Die lokale Erreichbarkeit des Objektes  $x$  ist definiert als:

$$lrd_{MinPts(x)} = 1 / \left\{ \frac{\sum_{x' \in N_{MinPts}(x)} rd_{MinPts}(x, x')}{|N_{MinPts}(x)|} \right\} . \quad (6)$$

Die Erkennung von Ausreißern erfordert einen Vergleich der Dichte unterschiedlicher Objektmengen. Eine dynamische Bestimmung der Dichte von Objektmengen wird über den Parameter  $MinPts$  erreicht. Dieser spezifiziert dabei eine minimale Anzahl an Objekten. Die lokale Erreichbarkeitsdichte eines Objektes  $x$  ist der Kehrwert der durchschnittlichen Erreichbarkeitsdistanz basierend auf den  $MinPts$ -nächsten Nachbarn von  $x$ . Duplikate führen dazu, dass die Dichte unendlich ist, sofern die Summe der Erreichbarkeitsdistanzen null ist.

Der LOF eines Objektes  $x$  ist definiert als:

$$LOF_{MinPts}(x) = \left\{ \frac{\sum_{x' \in N_{MinPts}(x)} \frac{lrd_{MinPts}(x')}{lrd_{MinPts}(x)}}{|N_{MinPts}(x)|} \right\} . \quad (7)$$



Zur Bestimmung des LOF kann neben der Euklidischen Distanz, die Manhattan Distanz (8) verwendet, hierbei werden die Differenzen nicht quadriert.

$$d(x, y) = \sqrt{\sum_i |x_i - y_i|}. \quad (8)$$

## 4 Analyse der Abrechnungsdaten

Im Folgenden werden die Abrechnungsdaten von Hausärzten analysiert. Aufgrund des Datenschutzes sind in der Analyse keine Sozialdaten aufgeführt. Eine Interpretation der Analyseergebnisse kann bei einem abweichenden Datenbestand unterschiedlich ausfallen und ist daher kritisch zu betrachten. Im Vordergrund steht das methodische Vorgehen. Im Folgenden werden Schritte der Datenanalyse beschrieben. Der Aufbau des Kapitels orientiert sich am KDD-Prozess gemäß Bild 1.

### 4.1 Auswahl

Die Zielsetzung der Analyse ist die Identifizierung von Fraud-Ärzten auf Basis von einheitlichen Betrugsmustern oder durch eine Abweichung vom Standardverhalten. Die Auswahl beschränkt sich auf die eingereichten Abrechnungsdaten der Hausärzte und die Information der Krankenkasse, welche Ärzte Abrechnungsbetrug begangen haben. Alle benötigten Informationen sind in einer Datenbank gespeichert. Ein Zugriff auf weitere Quellsysteme ist nicht notwendig. Der Geschäftszweck der genannten Rechenzentren ist die korrekte Abrechnungserstellung. Die Verarbeitung der personenbezogenen Daten ist als Mittel für die Erfüllung eigener Geschäftszwecke zulässig (§ 28 Absatz 2 BDSG). Tabelle 1 zeigt eine Übersicht der ausgewählten Daten. Für einen Versicherten liegt, je dokumentierter Leistung, ein Tupel vor. Fraud-Ärzte werden über den Binärwert Fraud-Arzt = 1 kenntlich gemacht.

Spaltenbezeichnung	Erklärung
ArztID	Primärschlüssel des Arztes
Fraud-Arzt	Binärwert zur Kennzeichnung von Fraud-Ärzten
VersichertenID	Primärschlüssel des Versicherten
Leistungsziffer	Primärschlüssel der Leistung
Leistungsbetrag	Honorarhöhe der Leistung
Leistungsdatum	Datum, an dem die Leistung erbracht wurde
Krankenkasse	Krankenkasse des Versicherten

Tabelle 1: Übersicht der ausgewählten Daten

### 4.2 Vorverarbeitung

Im Rahmen der Datenbereinigung wird geprüft, ob alle ArztID und VersichertenID einen numerischen Wert aufweisen. Zusätzlich kann über eine Moduloprüfung festgestellt werden, ob es sich um eine korrekte Versichertennummer handelt. Eine Harmonisierung der Daten ist nicht notwendig, da es sich um lediglich ein Quellsystem handelt. Es sollen nur Daten analysiert werden, die unter gleichen Voraussetzungen dokumentiert wurden. Aus diesem Grund wird die Auswahl der Abrechnungsdaten auf eine Krankenkasse und ein Quartal reduziert, da im Selektivvertrag krankenkassenspezifische Dokumentationskataloge vorliegen und diese sich quartalsweise ändern können.

### 4.3 Transformation

In dieser Phase werden die vorverarbeiteten Daten für die Clusteranalyse aggregiert. Die betrachteten Objekte sind die Ärzte. Die Variablen sind mit den Achsen des Vektorraums gleichzusetzen. Die metrischen Werte der Variablen werden gleichverteilt und kategorial einer Ordinalskala zugewiesen mit den Bereichen von beispielsweise 1 bis 20 (Beekmann 2003). Es werden folgende Variablen ausgewählt:

- Durchschnittlicher Fallwert der Versicherten.
- Durchschnittliche Anzahl an dokumentierten Leistungen pro Versicherter.
- Durchschnittliche Anzahl behandelte Versicherte pro Tag.

Die Auswahl der Variablen ist darin begründet, dass sie das Honorarvolumen eines Arztes am stärksten beeinflussen.

Tage, an denen ein Arzt keinen Versicherten abgerechnet hat, werden von der Berechnung der Durchschnittswerte nicht berücksichtigt, da nicht alle Betriebsstätten fünf Tage in der Woche geöffnet haben. Ferner wird der Datenbestand in Trainings- und Testdaten aufgeteilt (Beekmann 2003). Tabelle 2 zeigt das Schema der aggregierten Daten. Die Spalte „Datensatz“ dient als Unterscheidungsmerkmal. Für die Clusteranalysen umfassen die Trainings- und Testdaten jeweils 376 Ärzte und es sind jeweils drei bekannte Fraud-Ärzte enthalten.

ArztID	Fraud-Arzt	Durchschnittlicher Fallwert seiner Versicherten (1 - 20)	Durchschnittliche Anzahl an dokumentierten Leistungen pro Versicherter (1 - 20)	Durchschnittliche Anzahl behandelte Versicherte pro Tag (1 - 20)	Datensatz (Training oder Test)
--------	------------	--	---	--	--------------------------------

**Tabelle 2: Übersicht der aggregierten Daten**

### 4.4 Data-Mining

Im Folgenden werden die transformierten Daten gemäß des Vorgehens aus Kapitel 3 analysiert.

#### 4.4.1 Vorgehen 1

Ziel ist es, dass sich bekannte Fraud-Ärzte im selben Cluster befinden. Die Anzahl der Cluster wird, je Analyse, in Zwischenschritten hochgezählt, bis 38 Cluster erreicht sind. Die Auswahl der maximalen Clusteranzahl ist darauf zurückzuführen, dass im Datensatz circa zehn Prozent Fraud-Ärzte enthalten sind (Statistisches Bundesamt 2015). Bei einer gleichverteilten Anzahl von Ärzten zu Clustern entspricht dies 37,5 Cluster.

Der Lloyd-Algorithmus wird so lange wiederholt, bis sich die Mittelpunkte der Cluster nicht mehr verschieben. Die Ergebnismenge umfasst für jede Analyse eine Clusterzuordnung der Objekte. Das Training kommt zu dem Ergebnis, dass sich zehn Cluster am besten eignen, da die bekannten Fraud-Ärzte sich alle in Cluster neun befinden. Die Analyse mit zehn Clustern wird erneut auf Basis des Testdatensatzes durchgeführt und kommt zu dem Ergebnis, dass zwei von drei Fraud-Ärzten sich ebenfalls im selben Cluster befinden. Tabelle 3 zeigt die Ergebnisse der Analyse für die insgesamt sechs bekannten Fraud-Ärzte im Test- und Trainingsdatensatz.

ArztID	Cluster	Datenbestand
86	9	Trainingsbestand
226	9	Trainingsbestand
355	9	Trainingsbestand
604	6	Testdatenbestand
390	6	Testdatenbestand
438	2	Testdatenbestand

**Tabelle 3: Clusterzuweisung der Fraud-Ärzte aus dem Trainings- und Testdatenbestand**

#### 4.4.2 Vorgehen 2

Ziel der Analyse ist es, dass bekannten Fraud-Ärzten ein relativ hoher Local Outliner Factor zugewiesen wird. Die Distanzfunktionen Euklidische Distanz und die Manhattan Distanz werden verwendet, um unterschiedliche Gewichtungen für hohe Ausreiserwerte zu berücksichtigen. Für jede dieser Distanzfunktion werden jeweils 25 Analysen mit unterschiedlichen MinPts-Parametern durchgeführt. Der MinPts-Parameter beginnt bei zwei und wird für jede Analyse um zwei erhöht, bis 50 erreicht ist. Es werden somit je Distanzfunktion insgesamt 25 Analysen durchgeführt. Das Ranking der Ärzte wird anhand des LOF vorgenommen. Der Arzt mit dem höchsten LOF ist auf Rang eins. Ein Vergleich der Ergebnismengen kommt zu dem Ergebnis, dass die Manhattan Distanz und mit dem Parameter MinPts = 26 am besten geeignet ist, da die Fraud-Ärzte bei diesen Parametern im Durchschnitt den höchsten Rang aufweisen. Mit diesem Parameter wird die Analyse erneut auf Basis des Testdatensatzes durchgeführt. Tabelle 4 zeigt die Ergebnisse der Analyse im Trainings- und Testbestand mit dem Parameter MinPts = 26 und der Manhattan Distanz.

ArztID	Local Outliner Factor	Rang von 376 Ärzten pro Datenbestand	Datenbestand
86	1,240	56	Trainingsbestand
226	1,163	177	Trainingsbestand
355	1,867	2	Trainingsbestand
390	1,007	246	Testdatenbestand
438	1,311	4	Testdatenbestand
604	1,209	127	Testdatenbestand

**Tabelle 4: Local Outliner Factor der Fraud-Ärzte aus dem Trainings- und Testdatenbestand**

#### 4.5 Interpretation

Die Ergebnisse des „Vorgehen 1“ (vgl. Abschnitt 4.4.1) zeigen, dass der k-Means-Algorithmus trainiert wurde, sodass sich alle bekannten Fraud-Ärzte im Trainingsdatenbestand im selben Cluster befinden. Im Testdatensatz wurden zwei von insgesamt drei bekannten Fraud-Ärzten demselben Cluster zugewiesen. Es ist notwendig, dass sich Fraud-Ärzte aus unterschiedlichen Datenbeständen im selben Cluster befinden. Aus diesem Grund kann interpretiert werden, dass Fraud-Ärzte ein gemeinsames Dokumentationsverhalten ausweisen und weitere Ärzte in dem Fraud-Cluster mit einer erhöhten Wahrscheinlichkeit ebenfalls Fraud-Ärzte sind.

Die Ergebnisse von „Vorgehen 2“ (vgl. Abschnitt 4.4.2) zeigen, dass der Algorithmus zur Bestimmung des LOF trainiert werden konnte, sodass zwei von drei bekannten Fraud-Ärzten im Trainingsdatenbestand einen relativ hohen Local Outliner Factor ausweisen. Im Testdatenbestand hat einer von drei bekannten Fraud-Ärzten einen relativ hohen LOF. Aus diesem Grund weisen

dieser Auswertung Fraud-Ärzte kein Dokumentationsverhalten auf, welches sich signifikant von der nächstgelegenen Teilmenge an Ärzten unterscheidet. Würden die nächstgelegenen Objekte von Fraud-Ärzten ebenfalls Fraud-Ärzten sein, dann würden die Ergebnisse von „Vorgehen 2“ die Ergebnisse von „Vorgehen 1“ zusätzlich bestätigen.

## 5 Fazit

Es konnte festgestellt werden, dass zwischen Data-Mining und Fraud Detection ein Zusammenhang besteht. Kein Themengebiet umfasst das andere vollständig. Data-Mining beinhaltet lediglich den nicht-trivialen Teil von Fraud Detection. Ein vollständiges Fraud-Detection-System muss aus den nicht-trivialen und dem trivialen Prüfmechanismen bestehen. Eine mögliche triviale Prüfung wäre beispielsweise einen Schwellenwert für die maximale Anzahl an abgerechneten Versicherten pro Tag zu definieren.

Es wurden zwei Vorgehen für den Einsatz von Data-Mining-Clusterverfahren für Fraud Detection herausgearbeitet. In Abhängigkeit der Parameterwahl besteht die Möglichkeit, dass es sich um konkurrierende Vorgehen handelt. Sofern das „Vorgehen 1“ zu dem Ergebnis kommt, dass Fraud-Ärzte ein gemeinsames Dokumentationsverhalten aufweisen, wird es schwieriger, Fraud-Ärzte über dichte-basierte Ausreißerverfahren zu identifizieren. Die lokale Dichte eines Fraud-Arzt ist aufgrund der geringen Distanz zu weiteren Fraud-Ärzten hoch und nicht weiter auffällig. Zusammenfassend konnte gezeigt werden, dass Data-Mining-Clusterverfahren zur Aufdeckung von Abrechnungsbetrug genutzt werden können.

Im Vordergrund der Datenanalyse stand das methodische Vorgehen. Dies kann dazu führen, dass die Ergebnisse bei anderen Datenbeständen abweichen und die Interpretation anders ausfällt. Die Variablenauswahl für die Objekte der Clusteranalyse wurde darin begründet, dass diese das Honorarvolumen eines Arztes am stärksten beeinflussen. Eine Abhängigkeit zwischen Abrechnungsbetrug und Honorarvolumen wurde dabei implizit angenommen. Dies gilt es in weiteren Untersuchungen kritisch zu hinterfragen.

Die Abrechnungssysteme der Rechenzentren müssen zunehmend komplexere Plausibilitätsprüfungen durchführen können. Hierdurch besteht die Gefahr, dass die Anzahl an Fraud-Ärzten weiter steigt, da ein Abrechnungsbetrug aufgrund der steigenden Komplexität seltener auffällt. Aus diesem Grund ist zu erwarten, dass die Bedeutung von Data-Mining-Clusterverfahren in Zukunft für Fraud Detection weiter zunehmen wird.

In Kapitel 3 wurden zwei Vorgehen zur Aufdeckung von Fraud-Ärzten beschrieben. Abrechnungsbetrug ist ein branchenspezifisches Problem. Die in der Einleitung geschilderte Problemstellung lässt beispielsweise von der Gesundheits- auf die Versicherungsbranche übertragen werden. Hierbei ist es lediglich wichtig, dass die korrekten Eingangsvariablen identifiziert werden.

## 6 Literatur

AOK Bundesverband (2013) Zahlen und Fakten 2013. [http://www.aokbv.de/imperia/md/aokbv/info/nonperiodika/zuf\\_2013\\_ppt\\_final.pdf](http://www.aokbv.de/imperia/md/aokbv/info/nonperiodika/zuf_2013_ppt_final.pdf). Abgerufen am 31.10.2015

Association of Certified Fraud Examiners (2015) What is Fraud? <http://www.acfe.com/fraud-101.aspx>. Abgerufen am 31.10.2015

- Baesens B, Van Veronique V, Verbeke W (2015) Fraud Analytics Using Descriptive, Predictive, and Social Network Techniques: A Guide to Data Science for Fraud Detection. Hoboken
- Beekmann F (2003) Stichprobenbasierte Assoziationsanalyse im Rahmen des Knowledge Discovery in Databases. Wiesbaden
- Chamoni P, Beekmann F, Bley T (2010) Verfahren des Data Mining. In: Chamoni P; Gluchowski P (Hrsg.): Analytische Informationssysteme. Business Intelligence-Technologien und -Anwendungen. Berlin et al., S. 329 - 356.
- Breunig M, Kriegel HP, Ng R, Sander J (2000) LOF Identifying Density-Based Local Outliers. München und New York
- Bundeskriminalamt (2004) Bundeskriminalamt: Bundeslagebild Wirtschaftskriminalität. [http://www.bka.de/nn\\_193376/SharedDocs/Downloads/DE/Publikationen/JahresberichteUndLagebilder/Wirtschaftskriminalitaet/wirtschaftskriminalitaetBundeslagebild2004,templateId=raw,property=publicationFile.pdf/wirtschaftskriminalitaetBundeslagebild2004.pdf](http://www.bka.de/nn_193376/SharedDocs/Downloads/DE/Publikationen/JahresberichteUndLagebilder/Wirtschaftskriminalitaet/wirtschaftskriminalitaetBundeslagebild2004,templateId=raw,property=publicationFile.pdf/wirtschaftskriminalitaetBundeslagebild2004.pdf). Abgerufen am 31.10.2015
- Chen MS, Han J, Yu P (1996) Data-Mining: An overview from Database Perspective. IEEE Transactions on Knowledge and Data Engineering 6 (8): 866 - 883
- Cressey D (1973) Other People's Money. Toronto
- Frawley W, Piatetsky-Shapiro G, Matheus C (1992) Knowledge Discovery in Databases. AI Magazine 13: 57 - 70.
- Freiknecht J (2014) Big Data in der Praxis. Lösungen mit Hadoop, HBase und Hive. Daten speichern, aufbereiten, visualisieren. München
- Hansen HR, Mendig J, Neumann J (2015) Wirtschaftsinformatik. Berlin.
- IWD (2010) Zur Aufdeckung von Unregelmäßigkeiten im Rahmen der Abschlussprüfung. Düsseldorf
- Kassenvereinigung Bayern (2014) Honorarumsatz aus vertragsärztlicher Tätigkeit in Bayern. <https://www.kvb.de/fileadmin/kvb/dokumente/Praxis/Honorar/Analyse-Statistik/KVB-HON-Umsatzstatistik-4-2013.pdf>. Abgerufen am 31.10.2015.
- MacQueen J (1967) Some Methods for Classification and Analysis of MultiVariate Observations. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability (1): S. 281 - 297
- Nimwegen S (2009) Vermeidung und Aufdeckung von Fraud. Lohmar
- Petersohn H (2005) Data Mining. Verfahren Prozesse, Anwendungsarchitektur. München
- Pricewaterhouse Cooper (2012) Abrechnungsbetrug im Gesundheitswesen. <https://www.pwc.de/de/gesundheitswesen-und-pharma/assets/pwc-studie-zeigt-handlungsbedarf-fuer-krankenversicherungen.pdf>. Abgerufen am 31.10.2015
- Statistisches Bundesamt (2015) Pressemitteilung. 132/15. [https://www.destatis.de/DE/PresseService/Presse/Pressemitteilungen/2015/04/PD15\\_132\\_23611.pdf.pdf?\\_\\_blob=publicationFile](https://www.destatis.de/DE/PresseService/Presse/Pressemitteilungen/2015/04/PD15_132_23611.pdf.pdf?__blob=publicationFile). Abgerufen am 31.10.2015



# From sketching to modeling – Supporting end-users to elicit processes

Christopher Lentzsch<sup>1</sup> and Alexander Nolte<sup>1</sup>

<sup>1</sup> Ruhr-University of Bochum, Department Information and Technology Management,  
{lentzsch, nolte}@iaw.ruhr-uni-bochum.de

## Abstract

Using graphical models can be considered a common practice in modern organizations. These models serve a variety of purposes. They are used in software design, to analyze and design business processes and even as a basis for controlling. Especially in design models are also used by process stakeholders or future users of a system. Current modeling tools however are mainly built to the requirements of modeling experts and are thus hard to use for people that are not trained in using them. We developed and evaluated a system that allows users to draw shapes of model elements. These shapes are then automatically converted into elements of a modeling notation. Evaluating this system in comparison to current modeling tools that are based on regular screen controls we found that sketch-based interfaces potentially reduce barriers for lay modelers and provide better usability compared to traditional modeling interfaces. We also found that sketch based modeling potentially reduces the time required to model and identified means for improving our system.

## 1 Introduction

Using and creating graphical models is a common practice in research and practice alike. Models are e.g. used for software design, to analyze and design business processes or as a basis for controlling. During the course of this paper we will focus on the domain of business process modeling. This domain is special with respect to the distribution of knowledge about the processes that have to be modeled and modeling itself. People that work in processes or are affected by them (stakeholders) are familiar with those processes. They are however oftentimes not familiar with a modeling notation or with methods of model elicitation (Dumas et al. 2013). We consider these people to be **lay modelers**. Process analysts on the other hand know how to gather information about a process and visualize processes using a modeling notation. They are however oftentimes not knowledgeable about the process that has to be modeled. We consider these people to be **modeling experts**. This distribution of knowledge demands collaboration between lay modelers and modeling experts (Prilla et al. 2013). Collaboration between them usually takes place in workshops where lay modelers are supported by modeling experts in eliciting process models. Lay modelers verbally contribute their knowledge about a process which is then transferred into elements of a modeling notation by modeling experts (Herrmann 2009). Research however

indicates that limiting lay modelers to verbal contributions in a workshop setting reduces the effectiveness of modeling workshops (Bostrom et al. 1993) and potentially affects commitment to process changes in a negative way (den Hengst and de Vreede 2004). Recent research has shown that lay modelers are up to a certain extent able to understand, use and create models on their own (Hoppenbrouwers et al. 2010; Rittgen 2010; Nolte and Prilla 2013). In this context, tool support has been identified as one of the main barriers to engage lay modelers in creating models (Nolte and Prilla 2013). Existing tools are mainly built with the needs of modeling experts in mind and provide little to no guidance for lay modelers.

Hand drawn sketches can serve as a means to overcome the limitations of tool support. Studies conducted by Cherubini et al. (2007) indicate that software-developers frequently use hand drawn sketches in order to e.g. explain parts of their design to a co-worker. Results from that study indicate that software-developers use whiteboards or flip-charts for modeling because they are readily accessible and easy to use. There are however also limitations to using sketches on whiteboards or flip-charts for modeling especially with respect to re-using the resulting models afterwards. Since there is no need to use a specific modeling notation, every modeler potentially uses her own set of symbols. This makes it hard for others to understand and re-use the resulting models afterwards. Furthermore despite the opportunity to store models electronically e.g. by using electronic whiteboards or taking pictures of models, the resulting models are still hard to re-use afterwards because of the necessity for manual post-processing.

Taking the aforementioned limitations and advantages of model sketches into account we created a system that allows users to create models by drawing shapes of model elements which are then automatically transferred into elements of a modeling notation. We deliberately focused on model creation rather than on altering existing models as lay modelers are not expected to easily understand complex models and alter them on their own. Our system should allow any user – independent of their individual modeling knowledge – to elicit processes. During our evaluation we thus also focused on the following research question:

**RQ: How does using a sketch based modeling tool affect model creation compared to traditional tools?**

In what follows we will provide a review of different approaches of sketch-based modeling with respect to our scenario (section 2). Based on this review we developed a system (section 3) and conducted a study in which we compared it to existing tools (section 4). Results from this study are presented and analyzed in section 5. In the end we will provide a conclusion and indicate directions for future work (section 6).

## 2 Existing systems

There are multiple approaches that allow users to create process models using interactive surfaces. In what follows we will describe existing approaches and analyze them with respect to our scenario.

The first tool we found that uses sketches of model elements to create process models is SUMLOW which has been developed by Chen et al. (2008) for CASE-scenarios (Teichroew and Hershey 1976). It allows users to sketch activity and class diagrams using the unified modeling notation (UML, (OMG 2011a)). SUMLOW analyzes the user's strokes continuously in order to recognize the drawn elements. Drawn elements keep their sketchy appearance and are transferred into model elements when the user exports the created model. By not introducing formalization too early Chen



et al. (2008) aim to support exploring different ideas and solutions instead of focusing directly on the first idea that is drawn up. In our case this can however turn into a shortcoming as lay modelers cannot check whether their action led to the intended result.

Another approach is the tool Inkus that has been developed by Mangano and Sukaviriya (2010). Inkus takes the approach of SUMLOW even further and does not enforce the use of notation shapes. Inkus allows users to sketch diagrams freely, add tags to the sketched shapes afterwards to identify them as business process elements and convert them to a formalized business process model using BPMN (OMG 2011b). This approach offers many possibilities for using ad-hoc notations, adding comments or creating constructs which a modeler cannot express formally herself. In our scenario however the modelers are not by all means knowledgeable about a modeling notation. By not providing any guidance on how to draw correct object shapes users need to know the modeling notation very well. This approach thus might be more suitable for advanced users or converting models from one modeling notation to another.

The aforementioned approaches do not provide features for instant formalization which we identified above as being crucial for model understanding and reuse. They do not deal with operations such as moving elements around or deleting them. As our system however has to support deletion of elements and moving them around, we also reviewed other touch based modeling interfaces. In contrast to our approach these interfaces do not support sketching model elements. Approaches such as the one by Kolb et al. (2013) use gestures to open a menu from which BPMN elements can be selected. The shapes of the gestures themselves are thus not related to the actual element that will be created. These tools however support additional means to manipulate the model – e.g. move single or multiple elements, delete elements – and the current view. Defining gesture for these tasks is challenging because there is very little agreement on specific gestures for certain tasks as the gestures sets proposed by Frisch and Dachzelt (2013), Kolb et al. (2013), Wobbrock et al. (2007) and Wobbrock et al. (2009) are all considerably different. We reused the delete gesture proposed by Kolb et al. (2013) and Wobbrock et al. (2007). The delete gesture proposed by Frisch and Dachzelt (2013) and Wobbrock et al. (2009) – dragging the element out of the boundaries of the screen – was not applicable for us as we also wanted to support large screens. Additionally – as we are using sketches to create elements – we consequently implemented actions done on a whiteboard with a pen as single pointer unistroke gestures e.g. crossing out an element. Additional actions like moving elements or zooming are performed using multi-touch gesture (e.g. with two fingers) thus imitating a virtual tool change.

For pen-based electronic whiteboards Donaldson and Williamson (2005) built a system to create basic UML activity diagrams. The gestures directly correspond to shape of the formal elements and are formalized on-the-fly. Furthermore Donaldson and Williamson (2005) exploit the context of a gesture to reduce recognition problems and reduce the size of the gesture set.

However, none of the aforementioned approaches provides any form of user guidance in terms of correct usage of the modeling notation. They also do not mark sections of a model which violate the syntax of the modeling notation.

### 3 System

We created a system that allows users to create models using unistroke gestures corresponding to the actual element shape. Therefore the users do not need to learn a large set of gestures or hot-keys

but can simply apply their knowledge of element-shapes to perform actions. The shapes are directly transferred into the corresponding elements of the modeling notation.

Our system itself is web-based and can thus be used on any device that is capable of running a browser. A screenshot is provided in Figure 1. We decided for a system that recognizes gestures on the client level using the \$1-Recognizer that was originally jointly developed by the University of Washington and Microsoft Research for gesture recognition (Wobbrock et al. 2007). There is thus no necessity to send strokes to a server which does the recognition and sends the results back. This speeds up the recognition process. Furthermore the \$1-Recognizer is faster than comparable algorithms (e.g. Rubine's algorithm or Dynamic Time Warping (Wobbrock et al. 2007)). Furthermore the \$1-Recognizer is easily extensible, only has a very lightweight JavaScript code base and is thus well suited to be used within a web-based modeling tool. In our setting the \$1-Recognizer provided very accurate results. We pre-trained the recognizer with 40 sample sketches for each element. It is therefore not necessary to train the recognizer before each session and allows users to start right away.

For our system we chose SeeMe (Herrmann 2006) as a modeling notation. SeeMe only consists of a small set of elements and relations. It is thus easy to learn and understand. As stated before our goal is to support lay and expert modelers in the early stages process modeling to create simple or short models. We thus only used elements that are required to express basic processes. Features provided by our system are the possibility to create basic SeeMe elements such as roles, activities, entities and modifiers. Each of the elements can be directly created using the gestures shown in

Figure 1 below. Elements will be created at the location and size drawn by the user. Furthermore the system also allows users to create relations between elements. Relations can be created by simply drawing a single line from one element to another. The direction of the relation is based on the direction of the stroke gesture conducted by the user.

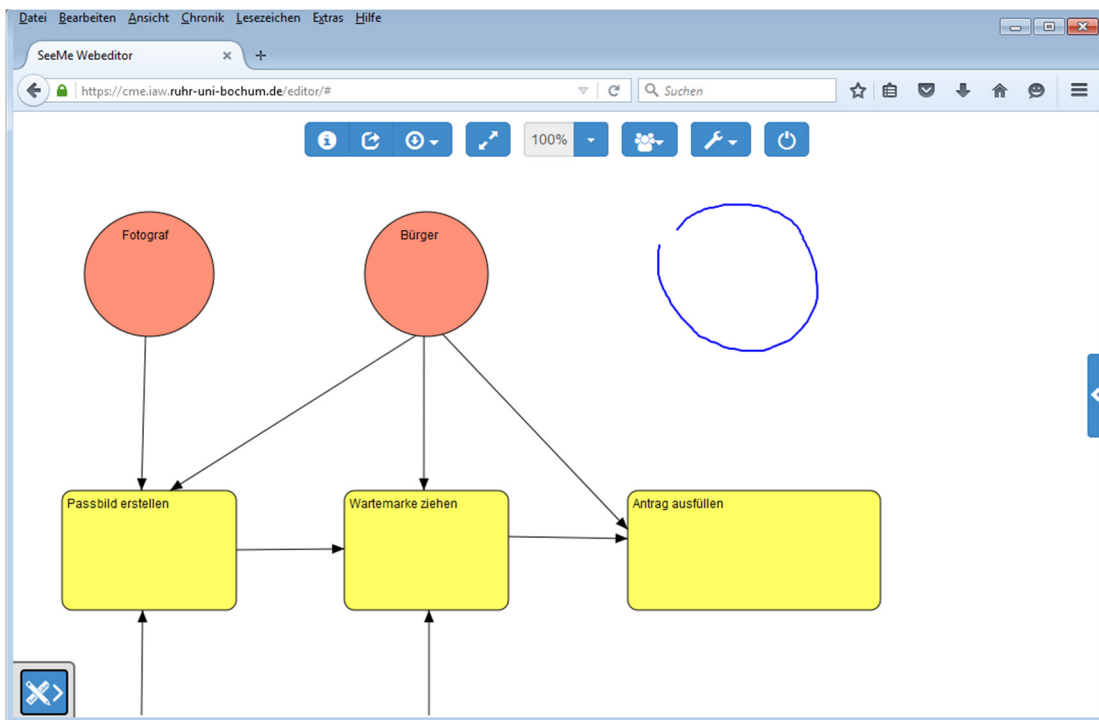
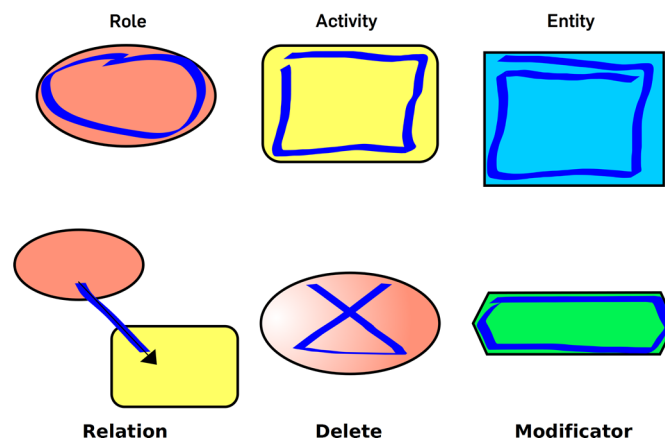


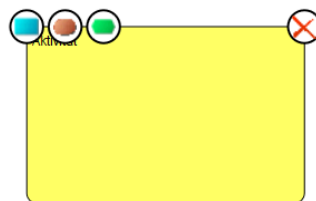
Figure 1 A Screenshot of our system

Figure 2 shows an overview of the gestures that we implemented in our system. The blue strokes represent the gestures that a user has to conduct in order to create the respective element. While roles and activities can easily be distinguished due to the difference in their shapes, the difference between activities and entities is not big enough to distinguish them automatically. We thus opted for adding a stroke to the gesture to create an entity as activities are generally used more in process models. This way the representation is still close to the actual element shape but is distinguishable by the user and the recognizer. Modifiers are created by drawing a hexagon. On creation they will be automatically attached to elements or relations if the starting point of the gesture is within or near such an element.



**Figure 2 Selection of gestures and their corresponding element or action**

We also provide means to alter elements after their creation. This is necessary as the system might fail to recognize the sketch correctly due to sloppy drawing or due to limitations of the algorithm. We thus provide a set of buttons that allows users to convert elements to the desired type or delete them completely (cf. Figure 3). The correction button set is attached to every newly created element or relation and offers users the possibility to delete an element after creation or transform it into an element of another type. Converted elements have the same size and location as the previously created element. After the conversion is done the buttons will be shown again to allow users to change the element type again or delete the element. The buttons are context sensitive and offer only element types that are different from the one that the algorithm detected.



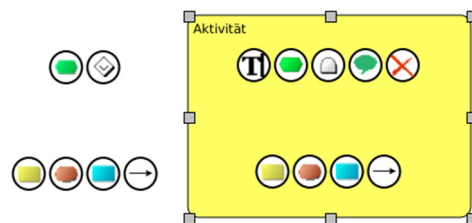
**Figure 3 Activity with the correction button set attached. The three buttons on the left offer means to transform the current element to either an entity, role or modifier. The rightmost button deletes the element.**

We also provided gestures to move elements, delete elements, move the canvas and adjust the zoom. These gestures were derived based on the set proposed by Wobbrock et al. (2009). In order to move an element two fingers are required and to pan the full hand or at least three fingers are needed

implementing the virtual tool change. However as these gestures are not within the main focus of the paper we will thus not describe them in detail.

## 4 Study

To answer our research question described in section 1 we conducted a comparative study. During the study we tested two conditions. The first condition was using the **sketch based interface** described before (condition: sketch). The second was an interface where elements can be created by touching the screen, selecting the corresponding element from an onscreen menu (cf. Figure 4) and then touching the desired position on the screen. This **button based interface** serves as a representative of traditional modeling tools (condition: button). Using both conditions alternately the participants were asked to redraw a process model that we created and printed before.



**Figure 4 Buttons on the canvas (left) and on an activity (right)**

Group A started modeling under condition sketch and group B under condition button first. Both process models described basic tasks likely to be known to the participants and were of comparable complexity and size. As a complexity metric we used connectivity as proposed by Mendling et al. (2007).

As participants for our study, we recruited students and graduates of applied computer science that are familiar with using multi-touch devices such as smartphones and tabletops. In total we had 8 participants aged between 20 and 35 years with 3 of them being lay modelers. As we did not conduct a controlled experiment but rather a study we used a mixed method approach (Creswell 2013) to answer our initial research question.

Before the study the user's general modeling expertise was captured with an adapted version of the questionnaire developed by Mendling et al. (2012) to distinguish lay modelers from modeling experts. The score is based on theoretical knowledge, prior modeling experience and intensity of modeling. Afterwards the participants were introduced to the respective interfaces and given the opportunity to ask questions regarding the model and the interface. The processes we used as tasks in our study were mostly linear and represented not very frequent but common tasks of everyday life like changing tires on a car or applying for an ID card.

After the introduction the participants were asked to start modeling the process. After they fulfilled one task they were asked to fill in the System Usability Scale questionnaire (Brooke 1996). The study ended with a semi-structured interview where the participants were asked to report on problems as well as their experience with the tools including their preference for one or the other condition.

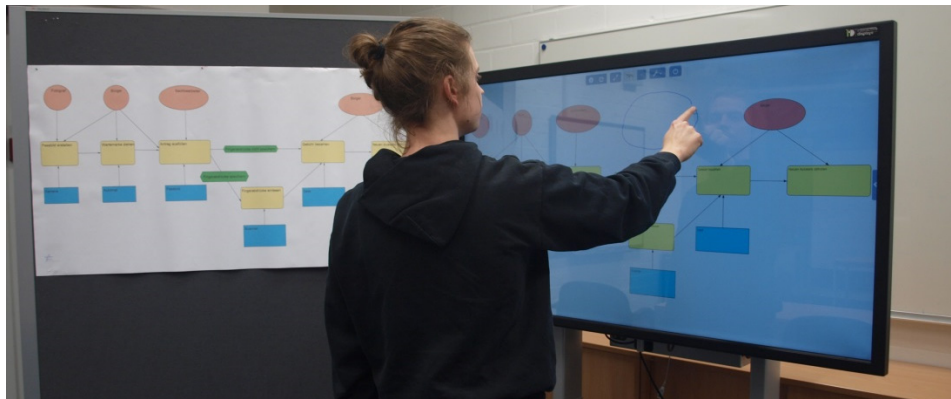


Figure 5 Setup of the study with the model that has to be reproduced on the left

## 5 Study results and analysis

Aiming at answering the question of how a sketch based modeling tool affects model creation compared to traditional tools (cf. RQ in section 1), we started by analyzing the time it took the participants to complete the task of reproducing a model using both the sketch and the button based system. We found that the time to complete the modeling task was shorter for the condition sketch every time compared to the condition button as indicated in Table 1. Furthermore every participant rated our system better than the button based approach, too.

We also compared the results between the groups that completed the task under condition sketch first (group A) to those that completed the task under condition button first (group B). We did not find any significant impact on the time needed to complete the task.

Participant	Group	Duration Sketching [s]	Duration Buttons [s]
<b>B</b>	A	198,820	224,755
<b>D</b>	A	298,886	345,027
<b>F</b>	A	228,822	330,080
<b>H</b>	A	286,012	330,168
<b>A</b>	B	338,012	398,341
<b>C</b>	B	326,071	432,646
<b>E</b>	B	226,053	354,030
<b>G</b>	B	389,439	397,042
<b>Mean</b>		<b>286,514</b>	<b>351,511</b>
<b>Standard deviation</b>		<b>65,069</b>	<b>63,109</b>

Table 1 Duration of the modeling tasks

Comparing group A's and B's performance combined reveals that all participants were faster when they used our sketching system. The very similar standard deviations for our system (SD:65,069) and the button based approach (SD:63,109) indicate that the given tasks were quite similar for the participants (cf. Table 1).

The analysis of the usability questionnaire shows similar results. All participants rated the sketching system higher than the button based approach resulting in a mean SUS score of 89,5 for the sketching system and 77 for the button based approach (cf. Table 2). Additionally all participants

stated in the concluding interview that they would prefer using sketches rather than the button based approach for the given modeling as well as their own modeling tasks.

<b>Participant</b>	<b>Group</b>	<b>Score Sketching</b>	<b>Score Buttons</b>
<b>B</b>	A	88	83
<b>D</b>	A	88	70
<b>F</b>	A	80	75
<b>H</b>	A	100	92,5
<b>A</b>	B	88	58
<b>C</b>	B	93	90
<b>E</b>	B	93	83
<b>G</b>	B	87,5	67,5
<b>Mean</b>		<b>89,438</b>	<b>77,188</b>
<b>Standard deviation</b>		<b>5,828</b>	<b>11,910</b>

**Table 2 Results of the System Usability Score questionnaire**

Comparing modeling experts to lay modelers we found little to no difference. While the fastest participant was a modeling experts, the difference was not that big (37,3s) on average. Comparing the performances with respect to the tools used reveals that lay modelers were on average 17,8s slower using the sketching system and 56,8s slower using the button based approach the modeling experts. Modeling experts rated both tools higher than the lay modelers.

In the concluding interview the participants mentioned that the button based interaction felt cumbersome as three interactions with the screen were required to create an element – a tap to open the button set, a tap to choose the element type and a drag to determine the element size and position. In contrast to that the participants referred to the interaction with the sketching system as being more ‘fluid’ and more ‘natural’. This was also the main argument the participants mentioned when we asked them to choose one system over the other. With respect to the button based system the participants had difficulties creating elements as they sometimes selected a tool out of the button set and forgot about their selection afterwards. Tapping on the screen subsequently resulted in tiny elements being created rather than the buttons being displayed again as expected by the user. This interrupted their actions and they needed to fix this modeling error and afterwards refocus on their original task. This behavior did not occur when the participants used the sketching system because it only offers one mode. Mangano et al. (2014) also found that the usage of single mode applications is beneficial as it uses the amount of distraction for users in informal design tasks.

Overall, we found that our system decreased the amount of time needed to complete the given task independently of the expertise of the user or the order of tool usage. Additionally we observed that all participants perceived the usability of the sketching system to be better compared to the button based approach. Having also found less errors when using the sketching system combined with the improved perceived usability also indicates that our system indeed lowers the barrier for lay modelers to create simple models as Nolte and Prilla (2013) identified. Using of sketches or iconic gestures thus seems to be a suitable way to provide users with a sufficient gesture set with very little amount of learning required.

Based on our analysis we also identified means to improve our system such as providing an additional easier to draw gesture for modifiers as a hexagon is not a shape users can draw easily.

The participants also stated that docking relations automatically to the edges of the elements was at times confusing as this on the one hand could change the meaning of the relation and on the other hand did not work as expected by the users. Our system attaches each relation to the nearest edge of its start and destination point, but the participants expected the relation to dock to the cut edge.

## 6 Conclusion and Outlook

We developed a system to support lay modelers and modeling experts during the early stage of modeling while paying particular attention to supporting lay modelers. In order to achieve this we combined features of sketch and gesture based modeling tools.

Despite the group of participants not being large enough for statistical analysis our results indicate that our system is faster and easier to use than traditional means of input. Additionally the users perceived the sketching system to perform considerably better compared to a button based system indicating that our choice of user interface is in sync with the users' expectations of sketch-based user interfaces. Therefore, our system is well suited for modeling experts as well as lay modelers.

In the future we are planning to extend the system with the possibility to add handwritten element labels or even use speech recognition as a means of input. Also worth considering are options to give hints on the correct usage of a notations syntax and warnings on common errors. This is especially crucial as we are planning to extend our set of sketches to other modeling notations such as the business process model and notation (BPMN (OMG 2011b) and the unified modeling language (UML (OMG 2011a)). We are also planning to conduct further studies with more users and different tasks in order to assess the feasibility of the approach as well as the tool for model sketching.

## 7 References

- Bostrom RP, Anson R, Clawson VK (1993) Group facilitation and group support systems. *Group support systems: New perspectives* 146–168.
- Brooke J (1996) SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189:4–7.
- Chen Q, Grundy J, Hosking J (2008) SUMLOW: early design-stage sketching of UML diagrams on an E-whiteboard. *Software: Practice and Experience* 38:961–994. doi: 10.1002/spe.856
- Cherubini M, Venolia G, DeLine R, Ko AJ (2007) Let's go to the whiteboard: how and why software developers use drawings. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, pp 557–566
- Creswell JW (2013) *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications
- den Hengst M, de Vreede GJD (2004) Collaborative business engineering: a decade of lessons from the field. *Journal of Management Information Systems* 20:85–114.
- Donaldson A, Williamson A (2005) Pen-based input of UML activity diagrams for business process modelling. In: *Proc HCI 2005 Workshop on Improving and Assessing Pen-based Input Techniques*, Edinburgh.
- Dumas M, La Rosa M, Mendling J, Reijers HA (2013) *Fundamentals of Business Process Management*. Springer

- Frisch M, Dachselt R (2013) Kombinierte Multi-Touch und Stift-Interaktion: Ein Gesten-Set zum Editieren von Diagrammen. In: Schlegel T (ed) *Multi-Touch*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 89–116
- Herrmann T (2009) Systems Design with the Socio-Technical Walkthrough. In: Whitworth B, de Moor A (eds) *Handbook of Research on Socio-Technical Design and Social Networking Systems*. Information Science Reference,
- Herrmann T (2006) SeeMe in a nutshell—the semi-structured, socio-technical Modeling Method. <http://www.imtm-iaw.rub.de/wp-content/uploads/sites/5/2011/09/Seeme-in-a-nutshell.pdf>. Accessed 29 Jun 2015
- Hoppenbrouwers S, Schotten B, Lucas P (2010) Towards Games for Knowledge Acquisition and Modeling. *International Journal of Gaming and Computer-Mediated Simulations*, Special issue on AI and Games 2:48–66.
- Kolb J, Rudner B, Reichert M (2013) Gesture-based Process Modeling Using Multi-Touch Devices. *International Journal of Information System Modeling and Design* 4:48–69.
- Mangano N, LaToza TD, Petre M, van der Hoek A (2014) Supporting Informal Design with Interactive Whiteboards. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, pp 331–340
- Mangano N, Sukaviriya N (2010) Inkus: a freehand method of creating business process models. In: *Proceedings of the Seventh Sketch-Based Interfaces and Modeling Symposium*. Eurographics Association, pp 143–150
- Mendling J, Reijers HA, Cardoso J (2007) What makes process models understandable? In: *Proceedings of the 5th international conference on Business process management*. pp 48–63
- Mendling J, Strembeck M, Recker J (2012) Factors of process model comprehension—Findings from a series of experiments.
- Nolte A, Prilla M (2013) Anyone can use Models: Potentials, Requirements and Support for Non-Expert Model Interaction. *International Journal of e-Collaboration (IJeC)* 9:45–60.
- OMG (2011a) Unified Modeling Language (UML), Version 2.4.1. <http://www.omg.org/spec/UML/2.4.1/>. Accessed 11 Jun 2015
- OMG (2011b) Business Process Model and Notation (BPMN), Version 2.0. <http://www.omg.org/spec/BPMN/2.0>. Accessed 11 Jun 2015
- Prilla M, Nolte A, Herrmann T, et al (2013) Collaborative Usage and Development of Models: State of the Art, Challenges and Opportunities. *International Journal of e-Collaboration Special issue on “Collaborative usage and development of models”* 9:1–16.
- Rittgen P (2010) Collaborative Modeling: Roles, Activities and Team Organization. *International Journal of Information System Modeling and Design (IJISMD)* 1:1–19.
- Teichroew D, Hershey EA (1976) PSL/PSA a computer-aided technique for structured documentation and analysis of information processing systems. In: *Proceedings of the 2nd international conference on Software engineering*. IEEE Computer Society Press, San Francisco, California, USA, p 2
- Wobbrock JO, Morris MR, Wilson AD (2009) User-defined gestures for surface computing. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, pp 1083–1092
- Wobbrock JO, Wilson AD, Li Y (2007) Gestures without libraries, toolkits or training: a \$! recognizer for user interface prototypes. In: *Proceedings of the 20th annual ACM symposium on User interface software and technology*. ACM, pp 159–168



# Cyber-Physische Produktionssysteme im Mittelstand: Herausforderungen zu Industrie 4.0 aus Sicht von KMU

Thomas Ludwig<sup>1</sup> und Volkmar Pipek<sup>1</sup>

<sup>1</sup> Universität Siegen, Institut für Wirtschaftsinformatik thomas.ludwig@uni-siegen.de,  
volkmar.pipek@uni-siegen.de

## Abstract

Das Schlagwort *Industrie 4.0* und die Vision smarterer Produktionssysteme erhalten immer stärkeren Einzug in die deutsche Industrielandschaft. Die technologische Ausgestaltung solcher smarten Produktionssysteme ist dabei vor allem unter dem ganzheitlichen Begriff der Cyber-physischen Produktionssysteme (CPPS) zusammengefasst. CPPS wird dabei verstanden als Konvergenz aus den Bereichen Maschinenbau, Sensorik und IT. Sowohl Industrie 4.0 als auch CPPS sind dabei aktuell stark an den Interessen der Großindustrie orientiert und vernachlässigen bei ihrer praktischen Ausgestaltung den Mittelstand – den größten industriellen Treiber Deutschlands – sowie dessen spezielle Anforderungen. Dieser Artikel führt in das Konzept von CPPS ein und präsentiert auf Basis eines Expertenworkshops, die Herausforderungen hinsichtlich CPPS und Industrie 4.0 aus Sicht von kleiner und mittlerer Unternehmen.

## 1 Einleitung

Die sogenannte vierte industrielle Revolution („Industrie 4.0“) zeichnet sich vornehmlich durch eine zunehmend komplexere Verbindung von Maschinen, Mitarbeitern, Materialien, Standorten und Unternehmen im Zeichen völlig veränderter Kommunikationsinfrastrukturen und Gewohnheiten aus. Diese Verbindungen werden zweifelsohne weitreichende Konsequenzen für die Produkte, das Verhältnis der Produkte zu Service- und Dienstleistungen, die eingesetzten Produktionsressourcen, Produktionsanlagen und den jeweiligen agierenden Mitarbeiter haben. Gleichzeitig steht die inner- und zwischenbetriebliche Organisation eines Unternehmens auf dem Weg zu einem dynamischen, echtzeitoptimierten und unternehmensübergreifenden Wertschöpfungsnetzwerk einigen Veränderungen entgegen.

Smarte Produktionssysteme und -anlagen gelten im Rahmen von Industrie 4.0 als (visionärer) Lösungsansatz, um den aktuellen Herausforderungen an die heutige und vor allem zukünftige Produktion entgegenzutreten, in dem die Lücke zwischen daten-, technologie- und prozessorientierten Produktionsgestaltungsformen geschlossen wird und dadurch eine Integration der virtuellen und realen Welt erreicht werden soll. Der Grundgedanke smarterer Produktionssysteme ist dabei möglichst „intelligente Produkte“ in den Produktionsprozessen einzusetzen, um die Abläufe jederzeit nachvollziehbar und damit das zugehörige Produktionssystem effizient und

flexibel zu gestalten. Solche intelligenten Produkte besitzen dabei neben ihren eigentlichen Funktionen auch informationstechnische Funktionen, mittels derer sie Daten speichern, kommunizieren und somit wieder bereitstellen können (Birkhahn 2007).

Technologische Treiber der smarten Produktionssysteme sind aktuell vor allem unter dem Begriff der Cyber-Physischen-Produktionssysteme (CPPS) bekannt. CPPS umfassen dabei verteilte, miteinander vernetzte, intelligente Produktionsanlagen, die eingebettete Systeme mit internet-basierten Funktechnologien vernetzen und Sensordaten aufnehmen sowie Aktoren steuern, um mit deren Hilfe die Material-, Güter-, und Informationsflüsse (oftmals autonom) zu regeln (Rajkumar et al. 2010). CPPS als Konvergenz aus den Bereichen Maschinenbau, Sensorik und IT soll dabei eine horizontale und vertikale Integration technischer und wertschöpfender Prozesse ermöglichen, mit dem Ziel, die gesamte übergreifende Produktion zu flexibilisieren.

Die deutsche Industrielandschaft ist durch ihre typische Mittelstandstruktur mit überwiegend kleinen und mittelständischen Unternehmen (KMU) geprägt. Beispielsweise haben in Nordrhein-Westfalen rund 95 Prozent der über 1.600 Maschinenbaubetriebe weniger als 500 Mitarbeiter und über zwei Drittel sogar weniger als 100 Beschäftigte (LeitmarktAgentur.NRW 2015). Der deutsche Mittelstand ist dabei geprägt durch innovative, maßgeschneiderte Produkte, einer hohen Spezialisierung und den Einsatz von Hochtechnologie – vor allem in den Bereichen Maschinen- und Anlagenbau, Automobil-Zulieferindustrie und Gebäudetechnik.

Scheint die Vision von Industrie 4.0 auf theoretischer und politischer Ebene bereits ausformuliert, ist die praktische Ausgestaltung von Industrie 4.0 und CPPS allerdings noch sehr vage formuliert und stark an den Interessen der Großindustrie orientiert. Betrachtet man Deutschland als mittelstandsgeprägten Produktionsstandort reicht es nicht aus, sich an der Großindustrie zu orientieren, sowie rein Technologie-zentrierte Entwicklungen ohne Betrachtung organisationaler, sozialer und wirtschaftlicher Wechselwirkungen in den Vordergrund zu rücken. Hat Industrie 4.0 bereits in einigen wenigen großen Industriekonzernen Einzug erhalten, gilt dies kaum für KMU (Wischmann et al. 2014). Wie Bundeswirtschaftsminister Sigmar Gabriel auf der Eröffnung des achten IT-Gipfels in Hamburg erklärte, stehe mittlerweile im Fokus, die gesamtwirtschaftlichen Folgen der Digitalisierung in den Blick zu nehmen, wobei es „besorgniserregend“ sei, dass 70 Prozent der mittelständischen Unternehmen in einer Umfrage zu Industrie 4.0 die Thematik Industrie 4.0 und Digitalisierung der Produktion sowie alle damit zusammenhängenden Potenziale und Risiken als nicht relevant für die Unternehmensstrategie ansehen (Greis 2014).

Aber warum nehmen KMU bei der Thematik Industrie 4.0 eine oftmals pessimistische Sichtweise ein? Der Mittelstand steht im Hinblick auf Konzepte und Technologien von Industrie 4.0 besonderen Herausforderungen gegenüber (Wischmann et al. 2014). So bewegen sich KMU oftmals in Nischenmärkten und produzieren Kleinstserien oder Einzelstücke nach speziellen Kundenanforderungen und mit einem historisch gewachsenen Mitarbeiter-bezogenen Fachwissen, worauf sich technische Konzepte von Industrie 4.0 – wie die von CPPS oder eines intelligenten Werkstücks – nicht ohne Weiteres adaptieren lassen. So ist der manuelle Fertigungsgrad von KMU in der Regel substantiell höher als der von Großunternehmen und das Arbeitsvermögen der Mitarbeiter – deren Wissen über Verfahren der Produktion – garantieren bereits über Jahrzehnte den Erfolg der Unternehmen innerhalb ihrer jeweiligen Nischenmärkte.

Obwohl Koch et al. (2014) vor allem aus technologischer und konzeptioneller Sicht zeigen, dass insbesondere für den Mittelstand Industrie 4.0 und die dadurch entstehende Möglichkeit, sich mit neuen Services zu differenzieren, einige Potenziale bietet, soll innerhalb dieses Artikels eine

kritische Perspektive eingenommen werden und es sollen die Herausforderungen des Mittelstandes hinsichtlich Industrie 4.0, sowie deren technologische Ansätze, aufgezeigt werden. Basierend auf der theoretischen Herleitung von CPPS als technisches Resultat einer kombinierten Daten-, Technologie- und Prozess-orientierten Produktionsgestaltung werden erste Ergebnisse in Form von zehn Thesen eines Workshops mit 15 Experten aus mittelständischen Industrieunternehmen hinsichtlich ihrer Bedenken gegenüber CPPS präsentiert. Diese ersten Ergebnisse sollen dazu verwendet werden Wissenschaftler, Unternehmensberatungen, Politiker, aber auch Praktiker in den relevanten Bereichen auf die Herausforderungen des Mittelstandes im Hinblick auf Industrie 4.0 und das mittelstandsspezifische Design von CPPS zu sensibilisieren.

## **2 Die Entwicklung von Cyber-Physischen Produktionssystemen**

Technologische Treiber von Industrie 4.0, welche vor allem Virtualisierung und Vernetzung sämtlicher (produktionsrelevanter) physischen Dinge über das Internet fokussieren, sind so genannte Cyber-Physische Systeme. Auf dem Weg zu Cyber-Physischen (Produktions-)Systemen lassen sich wissenschaftliche Ansätze industrieller Produktionsgestaltung hinsichtlich ihrer Fokussierung in die drei Bereiche daten-, technologie- und prozessorientierte Produktionsgestaltung unterscheiden (Birkhahn 2007).

Frühe datenorientierte Ansätze umfassen vor allem das Computer Integrated Manufacturing (CIM). Ausgehend von anfänglichen digitalen Unterstützungssystemen in der Konstruktion (CAE, CAD), Arbeitsplanung (CAP) und Fertigung (CAM, CAQ), beinhaltet das CIM durch Hinzunahme betriebswirtschaftlicher Produktionsplanung und -steuerungssysteme erstmalig ein integriertes Gesamtkonzept für die Informationsverarbeitung in der Prozess- und Fertigungsbranche. Der Grundgedanke von CIM besteht darin, die Konsistenz, Aktualität und Qualität von Unternehmensdaten mittels computergestützten Netzwerken zu erhöhen, um zu einer Verbesserung der Unternehmensprozesse beizutragen (Birkhahn 2007). Unter dem Schlagwort „digitale Fabrik“ sollen durch die Unterstützung geeigneter Software alle am Planungsprozess beteiligten Akteure auf ein digitales Modell und die hinterlegte Datenbasis zugreifen können.

Rein datenorientierte Ansätze, wie CIM oder darauf aufbauend die Konzeption einer digitalen Fabrik, führten allerdings nicht zu erhofften Erfolgen (Birkhahn 2007). Ziele, wie die Verkürzung der Durchlaufzeiten oder Personaleinsparung im Fertigungs- und Verwaltungsbereich, wurden aufgrund der oftmals Benutzer-unfreundlichen Bedienung und des impliziten Schulungsaufwands (Becker und Rosemann 1993) nicht in zufriedenstellender Weise erreicht (Büring 1997). Die Neu- und Weiterentwicklung von Technologien führte deshalb zu einer Reihe technologieorientierter Konzepte, wie die Smart Factory oder Smart Objects. Smart Objects sind hybride Produkte, die sich aus einer physischen und einer Daten verarbeitenden Komponente zusammensetzen (Fleisch 2001). Die Datenverarbeitung eines „intelligenten Dinges“ verbirgt sich im Hintergrund und wird vom Nutzer nicht offensichtlich wahrgenommen.

Im industriellen Kontext versucht das Konzept der Smart Factory die Lücke zwischen digitaler Planung und Realität (Lucke et al. 2008) zu schließen, wodurch ein transparentes, optimiertes Produktionsressourcen-Management realisiert werden soll, in dem hochdynamische Sensorinformationen in das kontextbezogene Umgebungsmodell integriert werden (Bauer et al. 2004). Rein technologieorientierte Ansätze werden jedoch vor einem historischen Hintergrund der digitalisierten Fertigung aufgrund der einseitigen Fokussierung kritisch betrachtet (Brödner 2015).

Die Flexibilität, mit denen Unternehmen auf dem globalen Markt reagieren müssen, führte zu prozessorientierten Ansätzen, die sich in frühen Philosophien der „Lean Production“ wiederfinden, bei der das Produktionsprogramm nicht wie bisher kapazitätsoptimiert ausgelegt wird, sondern nur noch das produziert wird, was der Kunde auch tatsächlich abnimmt (Womack et al. 1990). Lean Production bot zwar neue Lösungsansätze, jedoch waren die Veränderungen so weitreichend, dass sie vor allem in europäischen Produktionen einen Paradigmenwechsel bedeuten, da andere Rahmenbedingungen hinsichtlich Entlohnung, Produkthaftung, etc. gelten, als im japanischen Ursprungssystem (Birkhahn 2007). In jüngerer Zeit wurde Lean Production hin zu einem sozio-technischen Systems erweitert, das Unternehmen, Kunden und Zulieferer sowie entsprechende Feedbackschleifen integriert (Shah und Ward 2007).

Smarte Produktionssysteme gelten heutzutage als ein Lösungsansatz, um aktuelle Herausforderungen an die heutige Produktion zu meistern, indem die Vorteile der daten-, technologie- und prozessorientierten Produktionsgestaltung genutzt und kombiniert werden. Der Grundgedanke dabei ist es, sogenannte „wissensinkorporierte Objekte“ in Produktionsprozessen einzusetzen, um die Abläufe jederzeit nachvollziehbar und damit das zugehörige Produktionssystem sicher, effizient und flexibel zu gestalten. Solche Objekte stellen dabei Produktionsobjekte dar, die neben ihrer eigentlichen Funktion auch informationstechnische Funktionen besitzen, mittels derer sie Daten speichern und wieder bereitstellen können (Birkhahn 2007). Sie sind heute vor allem unter dem ganzheitlichen Begriff der Cyber-Physischen-Produktionssysteme bekannt und umfassen dabei verteilte, miteinander vernetzte, intelligente Produktionsmaschinen, die eingebettete Systeme mit Internet-basierten Funktechnologien verbinden und Sensordaten aufnehmen, mit denen sie Material-, Güter-, und Informationsflüsse durch Aktoren regeln (Rajkumar et al. 2010).

### **3 Herausforderungen des Mittelstandes hinsichtlich CPPS**

Die deutsche Industrie steht im internationalen Wettbewerb (vor allem im Low-Cost-Segment) der Marktübernahme insbesondere durch chinesische Hersteller gegenüber (Eisenhut et al. 2011). Für die mittleren und oberen Qualitäts- und Preissegmente bieten CPPS den deutschen Herstellern gute Chancen, um ihren Kunden durch zusätzliche Dienstleistungen einen wettbewerbsrelevanten Mehrwert zu bieten (Geisberger and Broy 2012; Bischoff et al. 2015). Deutsche Unternehmen sind dabei durch Selbstverpflichtungen wie ISO 9000ff oder gesetzliche Haftung für ihre Produkte und Prozesse gezwungen, stets ihre Produktionsprozesse selbst zu kontrollieren, zu dokumentieren und vor allem jederzeit zu beherrschen (Birkhahn 2007). Jedoch stellen die Komplexität entsprechender CPPS, der rasche technische Fortschritt sowie die Verknüpfung von Hard- und Software in dem Bereich der industriellen Produktion die kleinen und mittelständischen Unternehmen vor große Herausforderungen bei der Handhabung solcher komplexen Fertigungstechnologien (Ludwig et al. 2014).

Der aktuelle Stand der Produktionsgestaltung postuliert daher eine sehr an der Großindustrie orientierte und vor allem technokratische Top-Down-Sichtweise auf Industrie 4.0 (Wischmann et al. 2014). CPPS werden durch ein Zusammenwachsen von Informatik, Elektrotechnik und Maschinenbau sowie eine umfassende Ausstattung durch Sensorik bzw. Aktorik und deren Anbindung an Informations- und Kommunikationstechnologie (weiter-)entwickelt, ohne jedoch dabei (a) die individuellen Spezifika des Mittelstandes, (b) passende Organisationsstrukturen, sowie (c) geeignete wirtschaftliche Einführungsstrategien mitzuliefern.

CPPS existieren heute bereits zum großen Teil in Industrieunternehmen mit einem sehr hohen Automatisierungsgrad auf Basis weit entwickelter Automatisierungstechnik für

Produktionsanlagen (Stich et al. 2015). Dabei ist die Frage, wie sich die Konzepte von Industrie 4.0 und die CPPS als treibende Technologien so kombinieren und lokal oder global vernetzen lassen, dass sie den Anforderungen *aller* in der Wertschöpfungskette beteiligten Organisationen – auch den KMU – entsprechen und neue, effiziente und wirtschaftlich erfolgreiche Industrieprozesse und -systeme erschaffen. Es sind Konzepte und vor allem praxisorientierte Werkzeuge erforderlich, welche die mittelständischen Anwender in die Lage versetzen, mit der Entwicklung Schritt zu halten und effektiv sowie effizient zu arbeiten.

Es existieren bereits Studien, die auf Basis abgeschlossener oder noch laufender Projekte, retropektivisch die Herausforderungen, vor allem aber die zentralen Potenziale des Mittelstandes thematisieren (Wischmann et al. 2014; Bischoff et al. 2015). Um der ganzen Thematik offener gegenüber zu stehen und vor allem um die möglichen negative Auswirkungen von CPPS auf den Mittelstand zu eruieren, wurde ein Workshop mit 15 Experten, bestehend aus mittelständischen Geschäftsführern und Unternehmensberatern, mit Fokus auf deren Erwartungen und Bedenken an Industrie 4.0 und CPPS, durchgeführt. Dazu fand eingeleitet durch Impulsvorträge, ein Brainstorming und Diskussion aller Beteiligten statt. Die spätere Analyse der empirischen Daten versuchte, alle geäußerten Bedenken zu erheben. Der Workshop fand im Rahmen der *Siegener Mittelstandstagung* (<http://mittelstandstagung.de>) statt. Im Folgenden sind die zentralen Herausforderungen in zehn Thesen zusammengefasst, die aus kritischer Perspektive Industrie 4.0 und KMU betrachten und die es langfristig zu untersuchen gilt:

1. *Übertragbarkeit*: Es bestehen Bedenken, dass sich die in der Regel auf Großserien fokussierten CPPS nicht ohne weiteres auf KMU adaptieren lassen, da sich viele KMU in Nischenmärkten bewegen und oft Kleinserien oder Einzelstücke nach speziellen Kundenanforderungen herstellen. Gleichzeitig fehlt der erkennbare Grad notwendiger Industrie 4.0-Restrukturierung.
2. *Investitionssicherheit*: Es bestehen Bedenken, dass Investitionen in CPPS nicht nachhaltig am Markt bestehen oder über einen langen Zeitraum hinweg nicht gewartet werden, sowie zum anderen, dass sich CPPS in der eigenen Praxis nicht in einem wirtschaftlich vertretbaren Zeitrahmen amortisieren, da KMU einen weniger hohen Automatisierungsgrad besitzen.
3. *Substituierbarkeit*: Es bestehen Bedenken, inwieweit die Ideen und Konzepte der CPPS die eigene Wertschöpfungsposition als Zulieferer stärkt oder aber substituiert und in komplett automatisierten horizontalen Wertschöpfungsnetzwerken beliebig austauschbar macht.
4. *Mitarbeiterverlust*: Es bestehen Bedenken, dass Profile hochqualifizierter Mitarbeiter, als einer der wesentlichen Wissensgaranten von KMU, durch den Einsatz von CPPS verallgemeinert werden können und diese von Kooperationspartnern abgeworben werden können.
5. *Datensicherheit*: Es bestehen Bedenken, dass durch mangelnde Datensicherheit bei vernetzten CPPS die eigenen Unternehmensprozesse transparent werden und dadurch der Druck von Großkunden zunimmt, sowie Betriebsgeheimnisse an Konkurrenten verloren gehen können.
6. *Zuverlässigkeit des IT-Netzes*: Es bestehen Bedenken, dass die technische Vernetzung als grundlegendes Architekturelement von CPPS nicht zuverlässig funktionieren wird, was einen direkten Einfluss auf Geschäftsprozesse hat. Die Bedenken reichen von Lieferverzögerungen, über die vermehrte Produktion von Ausschuss, bis hin zu einem völligen Produktionsstillstand.
7. *Kontrollierbarkeit*: Es bestehen Bedenken, dass die eigenen Mitarbeiter das Verständnis über die zu einem Großteil autonom funktionierenden CPPS verlieren und bei Störfällen – trotz Selbstverpflichtungen wie beispielsweise ISO 9000ff – nicht mehr Herr ihrer Systeme sind.

8. *Ganzheitliche Betrachtung von CPPS*: Es bestehen Bedenken, dass die heterogene IT-Landschaft von Vertrieb und Lieferant nicht mit den CPPS zusammen funktioniert. Eine Identifikation notwendiger passgenauer Bausteine von CPPS, welche mit der heterogenen Systemlandschaft harmonisieren könnte, fällt schwer.
9. *Strategische Einführung*: Es bestehen Bedenken, dass eine Integration von CPPS und die Migration alter Daten nicht wirtschaftlich vertretbar durchgeführt werden kann. Es fehlen vor allem partielle Strategien, welche eine individualisierte Einführung von CPPS ermöglichen.
10. *CPPS-Zertifikate*: Es bestehen Bedenken, dass ähnlich zu den Normen des Qualitätsmanagements ISO 9001 auch CPPS-Zertifikate etabliert werden, welche KMU aus den Wertschöpfungsketten verdrängen, sollten sie diese nicht erfüllen und dass Großkunden den KMU den Einsatz und die Verwendung von CPPS diktieren werden.

Wie bereits oben genannte Studien aufzeigen, werden sich parallel zu den Herausforderungen auch Potenziale für den Mittelstand durch Industrie 4.0 und den Einsatz von CPPS ergeben (Bischoff et al. 2015). Mit diesem Beitrag wurde vor allem jedoch die kritische Perspektive von CPPS im Hinblick auf KMU fokussiert. Die obigen Thesen zu den Herausforderungen des Mittelstandes zeigen, dass sich die an der Großindustrie ausgerichteten Leitideen von Industrie 4.0 sowie die Konzepte und technologische Ausgestaltung in Form von CPPS nicht ohne weiteres geeignet auf den Mittelstand adaptieren lassen. Die Ergebnisse, wie beispielsweise die Bedenken hinsichtlich der Substituierbarkeit in der Wertschöpfungskette, der Mitarbeiterverlust oder die Befürchtung vor durch große Unternehmen determinierte CPPS-Zertifikate zeigen deutlich das Spannungsfeld auf, in welchem sich KMU bewegen. Soll die Leitvision von CPPS bei KMU daher vorangetrieben werden, besteht ein akuter Bedarf Industrie 4.0 im Kontext von KMU, deren Interessen und Stärken, sowie Bedenken und Risiken konkret in den Blickpunkt zu rücken. Denn speziell kleine und mittelständische Unternehmen stehen bei der Auswahl, Einführung und nachhaltigen Umsetzung von Industrie 4.0 Konzepten vor den Herausforderungen, zum einen geeignete und passende CPPS-Technologien für ihre oftmals sehr individuellen und historisch gewachsenen Problemstellungen herauszufinden, zum anderen die Wechselwirkungen zwischen Technologieauswahl und -einführung mit den bereits etablierten Organisations- und Prozessstrukturen, Mitarbeitern, sowie Unternehmenskulturen zu erfassen und dadurch den gesamten sozio-technischen Einführungs- und Umsetzungsprozess wirtschaftlich anzugehen.

Speziell diese Wechselwirkungen zwischen Technologie und Organisation erfordern kurzfristig einen Dialog über die Relevanz von CPPS im Kontext der KMU, sowie insbesondere über die Herausforderungen, die bezüglich ihres Designs und ihrer Integration in Produktionsprozesse bestehen (Poovendran 2010; Rajkumar et al. 2010). Dabei werden – angesichts der Komplexität und der sozio-technischen Ausrichtung solcher Systeme – Fragestellungen aus der Mensch-Maschine-Interaktion (z.B. Ludwig et al. 2014) sowie geeignete Einführungsstrategien eine immer größer werdende Bedeutung für deutsche KMU einnehmen.

Als nächste Schritte planen wir eine Fokussierung auf die sozio-technischen Spannungsfelder, welche bei der sozio-technischen Ausgestaltung und Einführung von CPPS entstehen werden. Auf Basis der zehn formulierten Thesen zu den aktuellen Bedenken werden wir neben den technischen Aspekten, vor allem auch die Auswirkungen auf den Mitarbeiter aus deren jeweiligen verschiedenen Perspektiven zu erfassen. Dazu soll neben der Unternehmenssicht insbesondere auch die Perspektive der Sozialpartner und Gewerkschaften herangezogen werden, um KMU nicht nur technisch, sondern vor allem sozio-technisch für Industrie 4.0 und CPPS aufzustellen.

## 4 Literatur

- Bauer M, Jendoubi L, Siemonheit O (2004) Smart Factory - Mobile Computing in Production Environments. In: Proceedings of the MobiSys, Workshop on Applications of Mobile Embedded Systems. Citeseer, pp 1–3
- Becker J, Rosemann M (1993) CIM und Logistik: Die effiziente Material- und Informationsflussgestaltung im Industrieunternehmen. Springer-Verlag Berlin Heidelberg
- Birkhahn C (2007) Smart Production Systems - intelligente Konzepte zur Gestaltung von Produktionssystemen.
- Bischoff J, Taphorn C, Wolter D, et al (2015) Studie „Erschließen der Potenziale der Anwendung von ‚Industrie 4.0‘ im Mittelstand“. Mülheim an der Ruhr
- Brödner P (2015) Industrie 4.0 und Big Data – wirklich ein neuer Technologieschub? In: Hirsch-Kreinsen H, et al (eds) Digitalisierung industrieller Arbeit. edition sigma 2015, Berlin.
- Büring E (1997) Anspruch und Realität computerintegrierter Produktionssysteme und Fertigungssteuerungskonzepte. Duncker & Humblot
- Eisenhut M, Lässig R, Liedl J (2011) Production Systems 2020.
- Fleisch E (2001) Betriebswirtschaftliche Perspektiven des Ubiquitous Computing. In: Buhl HU, Huther. A, Reitwiesener B (eds) Information Age Economy. Physica-Verlag, Heidelberg,
- Geisberger E, Broy M (2012) agenda CPS - Integrierte Forschungsagenda Cyber-Physical Systems.
- Greis F (2014) Gabriel fordert mehr IT-Begeisterung vom Mittelstand. [www.golem.de/news/it-gipfel-2014-gabriel-fordert-mehr-it-begeisterung-vom-mittelstand-1410-109985.html](http://www.golem.de/news/it-gipfel-2014-gabriel-fordert-mehr-it-begeisterung-vom-mittelstand-1410-109985.html). Accessed 16 Dec 2015
- Koch V, Geissbauer R, Kuge S, Schrauf S (2014) Industrie 4.0: Chancen und Herausforderungen der vierten industriellen Revolution.
- LeitmarktAgentur.NRW (2015) Produktion.NRW.
- Lucke D, Constantinescu C, Westkämper E (2008) Smart Factory - A Step towards the Next Generation of Manufacturing. Manuf Syst Technol New Front 115–118.
- Ludwig T, Stickel O, Boden A, Pipek V (2014) Towards Sociable Technologies: An Empirical Study on Designing Appropriation Infrastructures for 3D Printing. In: Designing Interactive Systems. Vancouver, Canada, pp 835–844
- Poovendran R (2010) Cyber-physical systems: Close encounters between two parallel worlds. In: Proceedings of the IEEE. pp 1363–1366
- Rajkumar R, Lee ILI, Sha LSL, Stankovic J (2010) Cyber-physical systems: The next computing revolution.
- Shah R, Ward PT (2007) Defining and developing measures of lean production. J Oper Manag 25:785–805.
- Stich V, Deindl M, Jordan F, et al (2015) Studie - Cyber Physical Systems in der Produktionspraxis.
- Wischmann S, Wangler L, Botthof A (2014) Industrie 4.0 - Volks- und betriebswirtschaftliche Faktoren für den Standort Deutschland. BMWi, Berlin
- Womack JP, Jones DT, Roos D (1990) The Machine that Changed the World: The Story of Lean Production.





# Negative externe Effekte bei der Nutzung mobiler Endgeräte – Zur Rolle der Privatsphäre Dritter im Entscheidungskalkül des Nutzers

Tina Morlok<sup>1</sup>, Christian Matt<sup>1</sup> und Thomas Hess<sup>1</sup>

<sup>1</sup> LMU München, Institut für Wirtschaftsinformatik und Neue Medien,  
[morlok, matt, thess]@bwl.lmu.de

## Abstract

Mobile Endgeräte mit Kamerafunktion erlauben es ihren Nutzern, bildliche Daten der Umgebung und somit auch von Dritten zu erfassen. Dies impliziert jedoch negative externe Effekte, da die Erfassung Dritter ein Eindringen in deren Privatsphäre darstellen kann. Bisherige Forschung hat sich vor allem darauf konzentriert, wie sich ein wahrgenommenes Eindringen in die eigene Privatsphäre auf das individuelle Technologienutzungsverhalten auswirkt. Da die Privatsphäre Dritter bei der Nutzung mobiler Endgeräte jedoch ebenfalls betroffen sein kann, ist es notwendig, deren Rolle im Entscheidungskalkül des Nutzers zu analysieren. Dieser Research-in-Progress Beitrag erweitert auf Basis der Communication Privacy Management Theory die bestehende Privatsphäre-Literatur um das Konzept der Privatsphäre Dritter. Der Beitrag liefert theoretische und methodische Grundlagen für die Untersuchung, wie sich die Wahrnehmung eines Eindringens in die Privatsphäre Dritter auf die individuelle Nutzungsentention der Technologienutzer auswirkt.

## 1 Motivation

Mobile Endgeräte mit Kamerafunktionen, wie Smartphones oder Action Cams, erleben in den letzten Jahren einen beispiellosen Zuwachs an Nutzungsraten (Huddleston 2015). Die Kamerafunktion erlaubt es den Nutzern, bildliche Daten der Umgebung in Form von Fotos bzw. Videos zu erfassen. Durch die gegebene Mobilität dieser Geräte erfolgt die Nutzung meist im unmittelbaren Umfeld von Dritten. Dadurch können diese Personen zum Subjekt der Aufnahme werden, was ein Eindringen in ihre Privatsphäre und somit negative externe Effekte zur Folge haben kann. Aufgrund der Verbreitung dieser Endgeräte und der Zunahme an Nutzungsszenarien kommt es immer häufiger zu Situationen, die sich negativ auf die Privatsphäre Dritter auswirken können. Der Verkaufstopp der Google Glass Anfang 2015 nach einer medial geführten Kontroverse (Barr 2015) zeigt, dass die Wahrnehmung eines Eindringens in die Privatsphäre Dritter bei der Nutzung mobiler Endgeräte sogar deren Diffusion behindern kann. Bislang unbekannt sind jedoch die Auswirkungen auf das Verhalten der Nutzer selbst.

Wie sich die Technologienutzungsintention bei einem gleichzeitig induzierten Eindringen in die Privatsphäre Dritter verändert stellt ein komplexes Phänomen dar, das bislang kaum erforscht wurde (Hoyle et al. 2014). Unklar ist, wie Nutzer subjektiv beurteilen, ob und wie stark sie bei der Nutzung der Kamerafunktion mobiler Endgeräte negative externe Effekte gegenüber Dritten verursachen und wie die Wahrnehmung darüber ihre eigene Technologienutzungsintention beeinflusst. Die bisherige Forschung hat den Einfluss unterschiedlicher Faktoren auf den individuellen Entscheidungsprozess zur Nutzung mobiler Endgeräte untersucht (Sutanto et al. 2013). Der Schwerpunkt lag jedoch darauf, wie sich eine vom Nutzer wahrgenommene Gefährdung der eigenen Privatsphäre auf dessen Entscheidungskalkül auswirkt (Smith et al. 2011). Um diese Forschungslücke zu schließen, zielt dieses geplante Forschungsvorhaben darauf ab, zwei aufeinander aufbauende Forschungsfragen zu beantworten:

*FF1: Welche Determinanten bestimmen, wie Nutzer mobiler Endgeräte das durch sie induzierte Eindringen in die Privatsphäre von Dritten wahrnehmen?*

*FF2: Wie beeinflusst die Wahrnehmung eines Eindringens in die Privatsphäre Dritter die eigene Technologienutzungsintention?*

Da zur Rolle der Privatsphäre Dritter im Entscheidungskalkül der Nutzer bisher kaum Forschung existiert, wird ein explorativer Forschungsansatz gewählt. Die Durchführung qualitativer Interviews auf Basis des Grounded Theory-Ansatzes (Glaser u. Strauss 1967) erlaubt es, die psychologischen Mechanismen im Rahmen des individuellen Entscheidungskalküls zu identifizieren und ein Verständnis über die nutzerseitige Wahrnehmung eines durch den Nutzer induzierten Eindringens in die Privatsphäre Dritter zu entwickeln. Die darauf aufbauende Online-Umfrage bietet zudem eine Validierung der identifizierten Determinanten und darüber, wie sich das wahrgenommene Eindringen auf die Technologienutzungsintention auswirkt.

Das geplante Forschungsvorhaben liefert einen Forschungsbeitrag, da es die theoretischen und methodischen Grundlagen für die Untersuchung der Rolle Privatsphäre Dritter im Entscheidungsprozess des Nutzers bereitstellt. Das Vorhaben erweitert das individuelle Entscheidungskalkül um einen bislang in der Forschung nicht berücksichtigten Faktor. Mit zunehmender Verbreitung mobiler Endgeräte mit Kamerafunktion sowie von Foto- und Video-Sharing-Plattformen, wie Instagram oder YouTube, wird diese Thematik auch für Anbieter immer wichtiger. Für sie ist es unabdingbar, Kenntnis über die Rolle der Privatsphäre bei der Technologienutzung zu erlangen, da mögliche Privatsphäverletzungen der Akzeptanz und Nutzung ihrer Technologien entgegenstehen können (Sutanto et al. 2013).

## **2 Theoretischer Hintergrund**

### **2.1 Communication Privacy Management Theory (CPMT)**

Individuen können bei der Nutzung mobiler Endgeräte negative externe Effekte verursachen, da sie auch die Privatsphäre Dritter gefährden. Daher ist es notwendig zu untersuchen, inwieweit Nutzer die Privatsphäre Dritter in ihrem Entscheidungskalkül berücksichtigen. Die Communication Privacy Management Theory (CPMT) eignet sich für den hier betrachteten Kontext, da sie ein konzeptionelles Verständnis für den Umgang mit der Privatsphäre anderer Individuen bereitstellt (Petronio 2002). Die Theorie erklärt mittels metaphorischer Grenzen wie Individuen mit ihrer und der Privatsphäre Dritter umgehen (Petronio 2010). Petronio (2002) unterscheidet dabei zwischen persönlichen und kollektiven Grenzen, die Individuen gleichzeitig managen. Persönliche Grenzen

beziehen sich auf die eigene Privatsphäre. Kollektive Grenzen beschreiben, wie Individuen mit der Privatsphäre anderer Personen umgehen. Petronio (2002) argumentiert, dass Individuen sich auch für die Privatsphäre anderer verantwortlich fühlen und sich als Miteigentümer sehen.

Die CPMT wurde ursprünglich entwickelt, um den Umgang mit Privatsphäre und persönlichen Informationen in zwischenmenschlichen Situationen, etwa in Familien (Petronio 2010), zu untersuchen. In den letzten Jahren wurde sie jedoch verstärkt auch im Online-Kontext (Blogs, e-Commerce) angewandt (Child et al. 2009). Obwohl die CPMT ein theoretisches Verständnis für den Umgang mit der Privatsphäre anderer Individuen liefert, haben sich die meisten Studien darauf beschränkt zu untersuchen, wie Individuen mit ihrer eigenen Privatsphäre umgehen. Auch wurde die CPMT und das Konzept der Privatsphäre Dritter bisher noch nicht im Kontext mobiler Endgeräte angewandt. Hier wird die Bedeutung kollektiver Grenzen jedoch besonders deutlich, da der Nutzer etwa bei der Verwendung der Kamerafunktion dieser Endgeräte in die Privatsphäre Dritter eindringen kann. Der Nutzer hat dabei weitestgehend die Kontrolle über die Erfassung Dritter und kann somit darüber entscheiden, ob er Miteigentümer von den Informationen Dritter wird. Es besteht folglich eine Machtasymmetrie zwischen dem Nutzer und Dritten. Daher ist es in diesem Kontext von besonderer Relevanz zu untersuchen, inwieweit der Nutzer die Privatsphäre Dritter in seinem Entscheidungskalkül berücksichtigt.

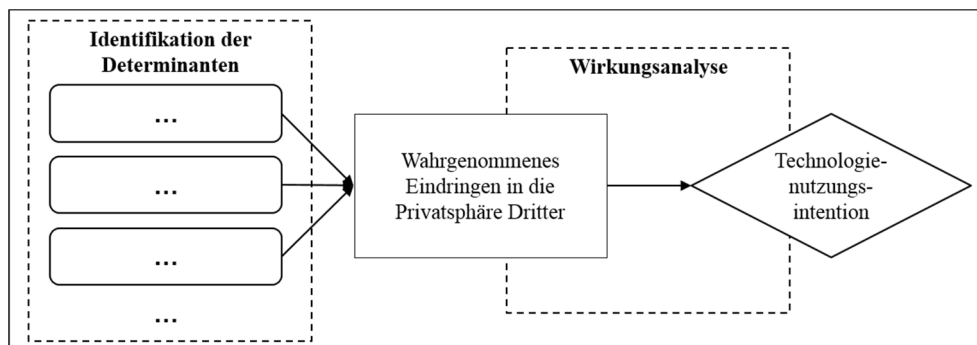
## **2.2 Das Eindringen in die Privatsphäre**

Das hier verwendete Konzept des Eindringens in die Privatsphäre basiert auf der Definition von Solove (2006), der das Eindringen als „die ungewollte Präsenz oder Aktivitäten einer anderen Person“ (Solove 2006, 550) beschreibt. Solove (2006) argumentiert, dass ein derartiges Eindringen die Aktivitäten des Betroffenen stört und dessen Routinen verändert. Das Konzept des Eindringens in die Privatsphäre basiert auf dem räumlichen bzw. physischen Verständnis von Privatsphäre, was als „der physische Zugang zu Individuen bzw. ihrer Umgebung und ihrem persönlichen Raum“ (Smith et al. 2011, 990) beschrieben werden kann. Durch die zunehmende Bedeutung von Informationstechnologien hat ergänzend dazu das Konzept der informationellen Privatsphäre in der Privatsphäre-Literatur verstärkt Beachtung gefunden. Die informationelle Privatsphäre beschreibt den „Zugang zu identifizierbaren persönlichen Informationen von Individuen“ (Smith et al. 2011, 990). Charakteristisch für den Kontext mobiler Endgeräte ist es, dass hier eine Kombination der beiden Konzepte von Privatsphäre auftritt.

Ebenso wie Gefährdungen der eigenen Privatsphäre, kann auch ein Eindringen in die Privatsphäre Dritter in unterschiedlichen Kontexten vorliegen. Beispielsweise verlangen einige App-Anbieter bei der Installation ihrer Dienste, dass Nutzer ihnen Zugriff auf die gespeicherten Kontakte gewähren, andernfalls kann die App nicht genutzt werden (Choi u. Jiang 2013). Auch in sozialen Netzwerken wird die Gefährdung der Privatsphäre Dritter besonders deutlich, da Nutzer hier Informationen über andere Personen preisgeben können, z.B. durch das Posten von Fotos oder Videos. Diese beiden exemplarischen Kontexte beschreiben jeweils mögliche Privatsphäreverletzungen beim Umgang mit Daten Dritter. Der in dem hier vorliegenden Beitrag betrachtete Kontext greift jedoch eine Stufe früher, nämlich bei der Entscheidung des Nutzers darüber, Daten anderer Individuen zu erfassen und dadurch deren physische und informationelle Privatsphäre zu gefährden. Unter Dritten werden hier (Un-)Beteiligte, den Nutzer umgebende Personen verstanden, die dem Nutzer (un-)bekannt sein können. In der vorliegenden Studie wird die Rolle der Privatsphäre Dritter bei der Nutzung mobiler Endgeräte untersucht. Diese Endgeräte umfassen u. A. Smartphones, Tablets, Wearables sowie Digital- und Action-Kameras (z.B. GoPro HERO). Die

Kamerafunktion dieser Geräte erlaubt es den Nutzern, bildliche Daten Dritter zu erfassen und somit in deren Privatsphäre einzudringen. Nutzer können Daten nicht nur digital erfassen und dadurch in die Privatsphäre Dritter eindringen, sondern aufgrund der Vernetzung der Geräte darüber hinaus Informationen Dritter online preisgeben. Neuere Entwicklungen, wie Wearables (z.B. Smart Watches, Datenbrillen) verstärken zudem die Gefährdung der Privatsphäre Dritter, da bildliche Aufnahmen mit diesen Geräten nahezu unbemerkt erfolgen können (z.B. Narrative Clip). Daher ist es notwendig, zu untersuchen, welche Rolle die Privatsphäre Dritter im Entscheidungskalkül des Nutzers mobiler Endgeräte spielt.

Einige wenige Feldexperimente in der Kerninformatik liefern erste Erkenntnisse für diesen Anwendungskontext (Hoyle et al. 2014). Jedoch fehlt bislang die theoretische Fundierung, um den kognitiven Entscheidungsprozess des Nutzers erklären zu können. Daher hat die hier vorliegende Studie insgesamt zwei Ziele: 1) die Identifikation der Determinanten, welche die nutzerseitige Wahrnehmung eines Eindringens in die Privatsphäre Dritter bestimmen und 2) den Einfluss des wahrgenommenen Eindringens in die Privatsphäre Dritter auf die Technologienutzungsintention zu erklären (vgl. Bild 1). Das hier betrachtete Konstrukt der Wahrnehmung eines Eindringens in die Privatsphäre Dritter beschreibt, inwieweit der Nutzer seine situationsspezifische Verwendung der Kamerafunktion als eine Verletzung der Privatsphäre Dritter bewertet. Die Technologienutzungsintention als abhängige Variable bildet ab, in welchem Ausmaß der Nutzer gewillt ist, die Kamerafunktion in einer spezifischen Situation zu verwenden.



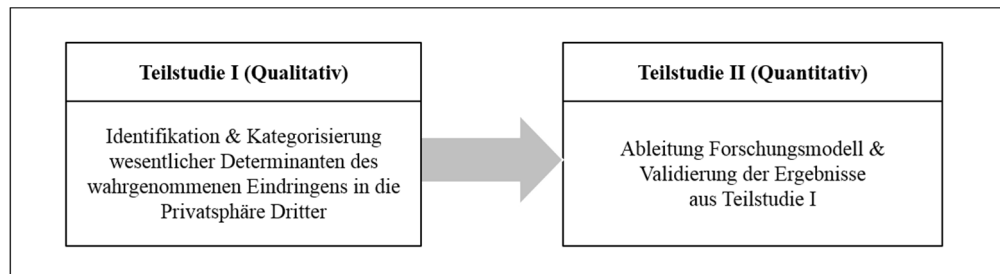
**Bild 1: Ziele der Studie**

### 3 Geplantes Forschungsvorhaben

#### 3.1 Aufbau und Ziele des Forschungsvorhabens

Das geplante Forschungsvorhaben (vgl. Bild 2) untergliedert sich in zwei Teilstudien. Die erste Teilstudie umfasst die Durchführung halbstrukturierter Interviews. Das Ziel ist hierbei die Bildung einer Kategorisierung derjenigen Determinanten, welche die nutzerseitige Wahrnehmung eines Eindringens in die Privatsphäre Dritter bestimmen. Die Durchführung der Interviews liefert die Grundlage für ein holistisches Verständnis darüber, anhand welcher Kriterien Nutzer das Eindringen in die Privatsphäre Dritter beurteilen. Einerseits werden wesentliche Determinanten der bestehenden Privatsphäre-Literatur abgefragt. So kann überprüft werden, ob diese auf den hier betrachteten Kontext und die Privatsphäre Dritter anwendbar sind. Andererseits ermöglicht der qualitative Ansatz die Identifikation zusätzlicher Determinanten, die speziell bei der Betrachtung der Privatsphäre Dritter im Kontext mobiler Endgeräte relevant sind. Basierend auf dieser Kategorisierung erfolgt eine quantitative Studie in Form einer Online-Umfrage. Diese hat zwei

Ziele: 1) Eine Validierung der identifizierten Determinanten und 2) eine Überprüfung, welchen Einfluss das Konstrukt des wahrgenommenen Eindringens in die Privatsphäre Dritter auf die Technologienutzungsintention hat.



**Bild 2: Vorgehensweise Forschungsvorhaben**

### 3.2 Qualitative Studie

Die qualitative Studie orientiert sich an dem Grounded Theory-Ansatz nach Strauss (Strauss u. Corbin 1990). Die Theorie ist für die Untersuchung dieses Kontexts besonders geeignet, da das Eindringen in die Privatsphäre Dritter durch die Nutzung mobiler Endgeräte bisher kaum untersucht wurde und daher Unklarheit über den kognitiven Entscheidungsprozess des Nutzers herrscht. Geplant sind halbstrukturierte Interviews, die auf jeweils ca. eine Stunde angesetzt und persönlich mit den Teilnehmern geführt werden. Damit sich die Teilnehmer in Nutzungsszenarien mit möglichen negativen externen Effekten gegenüber Dritten hineinversetzen können, werden Personen ausgewählt, die bereits seit mindestens einem Jahr aktiv ein Smartphone nutzen. Dieses Kriterium wurde gewählt, da mit zunehmender Dauer der aktiven Nutzung die Wahrscheinlichkeit steigt, dass sich die Teilnehmer bereits in derartigen Situationen befunden haben. Smartphone-Nutzer werden ausgewählt, da diese Klasse mobiler Endgeräte aktuell am stärksten verbreitet ist (McGrath 2015). Der zu Beginn entwickelte Basisfragebogen mit Kernfragen wird im Verlauf der parallel stattfindenden Datensammlung und -analyse immer weiter verfeinert. Die Interviewfragen sind darauf ausgerichtet zu verstehen, anhand welcher Determinanten die Teilnehmer als Nutzer mobiler Endgeräte das damit einhergehende Eindringen in die Privatsphäre Dritter beurteilen. Dabei werden auch Determinanten überprüft, die in der Privatsphäre-Literatur bereits einen Einfluss auf die Wahrnehmung eines Eindringens in die eigene Privatsphäre gezeigt haben. Hierbei werden kontextuelle als auch persönliche Faktoren berücksichtigt. Darunter fallen u. A. die Sensitivität der Informationen (Bansal u. Gefen 2010), soziale Normen (Xu et al. 2008) und der Ort der Aufnahme. Um die Interviews mit Hilfe der Analysesoftware ATLAS.ti auswerten zu können, werden diese mit einem Aufnahmegerät dokumentiert und anschließend vollständig transkribiert. Zur Analyse der Daten wird die Methodik des ständigen Vergleichens angewandt: Bereits erhobene Daten werden mit den entstehenden Kategorien und neu hinzukommenden Daten permanent verglichen bis „theoretische Sättigung“ erzielt worden ist (Glaser u. Strauss 1967).

### 3.3 Quantitative Studie

Im Rahmen der quantitativen Teilstudie findet eine Validierung der in den Interviews identifizierten Determinanten statt. Zudem wird der Einfluss des Konstrukts des wahrgenommenen Eindringens in die Privatsphäre Dritter auf die Technologienutzungsintention überprüft. Die Rekrutierung der Teilnehmer verläuft über Bekanntmachungen auf Facebook und über den E-Mail-Verteiler einer deutschen Universität. Filterfragen ermöglichen die Identifikation von Nutzern, die seit mindestens

einem Jahr aktiv Smartphones verwenden. Wo möglich, werden bestehende Skalen auf den hier betrachteten Anwendungskontext adaptiert, um Reliabilität und Validität dieser zu gewährleisten. Die Messung des Konstrukts *wahrgenommenes Eindringen in die Privatsphäre Dritter* basiert auf der von Xu et al. (2008) entwickelten Skala. Zudem werden Kontrollvariablen, wie Soziodemographika und Innovationsfreude (Xu et al. 2011), berücksichtigt. Die Validierung der Instrumente und des Forschungsmodells beruht auf einem Ansatz der Strukturgleichungsmodellierung (Partial Least Squares, PLS). So kann unter anderem auf Basis einer konfirmatorischen Faktorenanalyse eine Überprüfung der Validität der Multi-Item-Skalen erfolgen.

#### **4 Erwarteter Theorie- und Praxisbeitrag**

Obwohl frühere Studien bereits das individuelle Entscheidungskalkül bei der Nutzung mobiler Endgeräte untersucht haben, wurde die Rolle der Privatsphäre Dritter dabei bisher nicht berücksichtigt. Um die bestehende Forschungslücke zu schließen, wird im Rahmen der geplanten Studie eine Kategorisierung der Determinanten zur nutzerseitigen Wahrnehmung eines Eindringens in die Privatsphäre Dritter hergeleitet und mittels einer quantitativen Studie validiert. Die Studie liefert aufgrund des qualitativen und quantitativen Ansatzes ein ganzheitliches Verständnis zur individuellen Technologienutzungsintention. Die in der qualitativen Teilstudie geplante Kategorisierung beschreibt die wesentlichen Determinanten, welche die Wahrnehmung eines Eindringens bestimmen. Die darauf aufbauende Online-Umfrage liefert zudem eine Validierung dieser Determinanten sowie des Zusammenhangs zwischen dem wahrgenommenen Eindringen in die Privatsphäre Dritter und der Technologienutzungsintention der Anwender.

Die geplante Studie erweitert die bestehende Privatsphäre-Literatur, da sie Erklärungsansätze für die Rolle der Privatsphäre Dritter im Entscheidungsprozess des Nutzers bereitstellt. Ein Faktor, der in der bisherigen Literatur nicht untersucht wurde. Die Kategorisierung der Determinanten ermöglicht es, die Wahrnehmung einer Gefährdung der Privatsphäre Dritter im Entscheidungskalkül des Nutzers zu integrieren. Mit steigender Verbreitung mobiler Endgeräte mit Kamerafunktion wird die Analyse von Nutzungsszenarien mit möglichen negativen externen Effekten gegenüber Dritten immer wichtiger. Es ist daher wesentlich, den kognitiven Entscheidungsprozess der Nutzer zu verstehen und Theorie in diesem Forschungsbereich zu entwickeln. Auch für die Akzeptanz von neuen Technologien kann die Privatsphäre Dritter von maßgeblicher Bedeutung sein, wie es der Misserfolg einer ersten Markteinführung von Google Glass verdeutlicht (Barr 2015). Dieses Beispiel zeigt, dass eine bei der Technologienutzung wahrgenommene Gefährdung der Privatsphäre Dritter dazu führen kann, dass Individuen sich gegen einen Kauf bzw. eine aktive Nutzung dieser Technologien entscheiden. Unternehmen sollten sich darüber bewusst sein, dass nicht nur die eigene Privatsphäre, sondern auch die von Dritten für den Erfolg einer Technologie relevant sein kann. Diese Studie hilft Unternehmen zu verstehen, welche Privatsphäreaspekte den Nutzern bei der Verwendung mobiler Endgeräten mit Kamerafunktion wichtig sind. Wenn Anbieter hierfür ein Verständnis entwickeln, können sie künftig adäquat reagieren und eventuell notwendige Anpassungen umsetzen, sodass es zur Akzeptanz auf Seiten der Nutzer und Dritten kommt.

#### **5 Ausblick**

Aufgrund der zunehmenden Diffusion mobiler Endgeräte und somit möglichen Auswirkungen auf die Privatsphäre Dritter bedarf es künftig weiterer Forschung in diesem Feld. Neben der

Betrachtung der Technologienutzungsintention sollten weitere Untersuchungen zum tatsächlichen Nutzungsverhalten erfolgen, z.B. mit Hilfe von Feldexperimenten. Neben der Analyse des Nutzerverhaltens bei der Datenerfassung von Dritten bietet darauf aufbauend der individuelle Entscheidungsprozess zum Umgang mit bereits erfassten Daten Dritter ein interessantes Forschungsfeld.

## 6 Literatur

- Bansal G, Gefen D (2010) The Impact of Personal Dispositions on Information Sensitivity, Privacy Concern and Trust in Disclosing Health Information Online. *Decision Support Systems* 49(2):138-150
- Barr A (2015) Google Glass Gets a New Direction. *The Wall Street Journal*. <http://www.wsj.com/articles/google-makes-changes-to-its-glass-project-1421343901>. Abgerufen am 22.07.2015
- Child JT, Pearson JC, Petronio S (2009) Blogging, Communication, and Privacy Management: Development of the Blogging Privacy Management Measure. *Journal of the American Society for Information Science and Technology* 60(10):2079-2094
- Choi CFB, Jiang ZJ (2013) Trading Friendship for Value: An Investigation of Collective Privacy Concerns in Social Application Usage. *Proceedings of the 35th International Conference on Information Systems (ICIS)*. Milan, Italy
- Glaser BG, Strauss A (1967) *The Discovery of Grounded Theory: Strategies for Qualitative Inquiry*. Aldin, Chicago
- Hoyle R, Templeman R, Armes S, Anthony D, Crandall D, Kapadia A (2014) Privacy Behaviors of Lifeloggers Using Wearable Cameras. *ACM International Joint Conference on Pervasive and Ubiquitous Computing*. Seattle, USA
- Huddleston T (2015) Gopro's Revenue Goes Wild with Strong Overseas Sales. *Fortune*. <http://fortune.com/2015/04/28/gopro-quarter-virtual-reality/>. Abgerufen am 06.08.2015
- McGrath F (2015) Gwi Device: The Latest Trends in Smartphone, Tablet and Wearables Usage. *globalwebindex*. <https://www.globalwebindex.net/blog/gwi-device-the-latest-trends-in-smartphone-tablet-and-wearables-usage>. Abgerufen am 10.08.2015
- Petronio S (2002) *Boundaries of Privacy: Dialectics and Disclosure*. State University of New York Press, Albany
- Petronio S (2010) Communication Privacy Management Theory: What Do We Know About Family Privacy Regulation? *Journal of Family Theory & Review* 2(3):175-196
- Smith HJ, Dinev T, Xu H (2011) Information Privacy Research: An Interdisciplinary Review. *MIS Quarterly* 35(4):989-1016
- Solove DJ (2006) A Taxonomy of Privacy. *University of Pennsylvania Law Review* 154(3):477-564
- Strauss A, Corbin JM (1990) *Basics of Qualitative Research*. Sage, Newbury Park
- Sutanto J, Palme E, Tan C-H, Phang CW (2013) Addressing the Personalization-Privacy Paradox: An Empirical Assessment from a Field Experiment on Smartphone Users. *MIS Quarterly* 37(4):1141-1164
- Xu H, Dinev T, Smith HJ, Hart P (2008) Examining the Formation of Individual's Privacy Concerns: Toward an Integrative View. *Proceedings of the 29th International Conference on Information Systems*. Paris, France
- Xu H, Luo XR, Carroll JM, Rosson MB (2011) The Personalization Privacy Paradox: An Exploratory Study of Decision Making Process for Location-Aware Marketing. *Decision Support Systems* 51(1):42-52





# Decision Support and Data Visualization for Liner Shipping Fleet Repositioning

Daniel Müller<sup>1</sup> and Kevin Tierney<sup>1</sup>

<sup>1</sup> Universität Paderborn, DS&OR Lab, {Mueller | Tierney}@dsor.de

## Abstract

Liner carriers regularly move vessels between routes in their networks several times a year in a process called fleet repositioning. There are currently no decision support systems to allow repositioning coordinators to take advantage of recent algorithmic advances in creating repositioning plans. Furthermore, no study has addressed how to visualize repositioning plans and liner shipping services in an accessible manner. Displaying information such as cargo flows and interactions between vessels is a complex task due to the interactions of container demands and long time scales. Integrating visualization techniques for the repositioning problem in decision support systems gives liner shipping companies the opportunity to save money through better fleet utilization and cargo throughput, as well as reduce their environmental impact by using less fuel.

## 1 Introduction

Seaborne trade plays a major role in the world economy and is responsible for the transportation of about 9.6 billion tons of goods per year. In particular, the transport of containerized goods grew by 4.6 percent in 2013 and the total volume reached 1.5 billion tons or 160 million 20-foot equivalent units (TEU). Except for a short pullback during the financial crisis of 2008, the liner trade has been growing for decades, meaning that more and more goods are transported by sea (UNCTAD RMT 2014). In order to participate in the seaborne trade, liner carriers have to make large investments and are confronted with high daily operating costs (Fagerholt 2004). Sophisticated planning of a carrier's operations is a necessity in order to compete in this growing market. However, throughout the liner shipping industry the planning of networks, including route construction and vessel movements, is still performed manually. The emergence of new mathematical models and algorithms offers liner carriers the opportunity to improve their processes with decision support systems (DSS).

Containerized goods are transported on cyclical routes called services. Liner shipping companies operate a number of these services in order to connect different trade regions within their network. They are operated with a certain periodicity (usually weekly or biweekly) such that ports of a service are visited at a fixed time each period (*port call*). Furthermore, services consist of multiple slots, where each slot is associated with a specific vessel and a fixed schedule for the port calls. Depending on the periodicity of the service, there are fixed time intervals between the schedules of the slots.

Liner carriers must adjust their service networks on a regular basis due to trends in the world economy and seasonal variations in cargo volumes. Carriers add, remove and modify services to adjust to these changes. To carry out these modifications to the network, vessels must be moved, or repositioned, to other services. Repositioning coordinators generate an extensive plan for each repositioning of one or more vessels with the goal of minimizing the costs for moving the vessel(s) while maximizing the revenue earned from transporting customer cargo. This problem is referred to as the liner shipping fleet repositioning problem (LSFRP) (Tierney 2015).

Despite algorithmic advances in recent years to support repositioning coordinators, these advances are not yet accessible for non-researchers. To the best of our knowledge, there is no DSS for the LSFRP in use at any carrier. Repositioning coordinators are therefore still using spreadsheets (or other manual mechanisms) to create their plans. The process of creating a repositioning plan can therefore take up to a couple of days (Tierney 2015), which is significantly more than the several minutes required by modern heuristics (Tierney 2015). Although systems exist for other problems in the area of liner shipping (e.g. ship routing and scheduling), they lack specific requirements for the repositioning problem and thus only offer limited guidance in creating a DSS for the LSFRP.

A DSS that combines access to algorithmic approaches for the LSFRP as well as a comprehensible and usable approach to display relevant data would create enormous benefits for liner shipping companies. As fleet repositioning coordinators have lots of expert knowledge about the problem, they should be able to influence the behavior of such a system to an extent where they can make ad-hoc adjustments of models, constraints or other relevant aspects of the system.

In this paper, we therefore present an interactive DSS for the LSFRP. We also discuss various visualization techniques to display relevant data of the liner shipping fleet repositioning and general aspects of liner shipping. We believe the visualization strategies we present have a wide application within decision support systems in the maritime sector beyond fleet repositioning.

This paper is structured as follows. Section 2 contains an overview of the literature about the LSFRP followed by a description of liner shipping fleet repositioning in Section 3. In Section 4, we present options for visualizing the fleet repositioning problem. Finally, in Section 5, we conclude this paper and give an overview of our future plans for this research topic.

## 2 Liner shipping fleet repositioning

Whenever a vessel is reassigned to a different service, a repositioning coordinator must plan the actions of the vessel during the repositioning. He or she plans the repositioning within a horizon bounded by the earliest time the vessel may stop sailing according to its original schedule and the time it must begin operations on its new service. The time it starts the repositioning process is called the *phase-out*, and the time it ends the repositioning is called the *phase-in*.

Each repositioning of a vessel is associated with costs due to lost revenue and bunker consumption. The goal of fleet repositioning is to maximize the profit of carrying customer's cargo and moving empty containers in the network. A poor repositioning can result in a significant disruption to the carrying of customer's cargo, thus it is important for models to take cargo flows into account.

A number of activities exist for moving vessels between services cheaply. Slow-steaming, i.e., sailing a vessel below its design speed, can be used to lower sailing costs. Another option is to use so-called sail-on-service opportunities to sail between the original service and the goal service. Here, the repositioning vessel replaces another vessel (*on-service vessel*) in a specified service. The

on-service vessel can then be laid up or chartered to another company. In the repositioning plan, the coordinator can also add or remove ports in order to maximize the utilization of the vessel while regarding the fixed time period between phase-out and phase-in. A more detailed description of the liner shipping fleet repositioning problem can be found in Tierney (2015).

Due to the long time scales in liner shipping and the interactions between services and ports, fleet repositioning is a complex problem. Nevertheless, repositioning coordinators are still using spreadsheet software to manage their operations, which limits their abilities to visualize and comprehend complicated repositioning procedures, even when such plans would save money.

### 3 Literature Review

Despite the fact that the shipping industry plays a major role in the world economy, the number of DSSs in the literature for this industry is surprisingly few. In the area of bulk shipping, Fisher and Rosenwein (1989) describe a system that finds optimal fleet schedules for bulk cargo pickup and delivery with a dual algorithm. This system uses Gantt-charts, showing the individual schedules of the vessels, and maps for the itinerary of a selected vessel. Another system by Bausch et al. (1998) for the transport scheduling of bulk products uses spreadsheets and also Gantt diagrams to display and manage the proposed schedules. One of the more recent DSSs is TurboRouter by Fagerholt and Lindstad (2007). This system focuses on scheduling and routing decisions of the liner shipping companies. TurboRouter supports manual planning and optimization algorithms for automatic planning. Fagerholt and Lindstad (2007) report a significant improvement of the quality of the schedules and the time required for the planning process of the companies using their system.

Similar to the small amount of literature about DSSs in liner shipping, there are only a few publications that focus on the LSFRP. Even in their extensive list of relevant publications in the field of liner shipping, Christiansen et al. (2013) do not refer to the LSFRP, showing that this area of research has been explored only recently. The problem itself was first introduced by Tierney and Jensen (2012) and Tierney et al. (2012), in which they present the first mathematical approaches for solving the LSFRP. Furthermore, Tierney et al. (2015) develop a simulated annealing algorithm with cargo flows, in order to scale to large, real-world problems. The approach finds higher profits compared to a reference scenario from industry. The runtime of this algorithm is small enough such that the authors claim it is suitable for a decision support system.

There is some significant work for other problems in the field of liner shipping like the network design problem (NDP), as described by Agarwal and Ergun (2008) and Brouer et al. (2014), and the vessel schedule recovery problem (VSRP), as described by Brouer et al. (2013). The goal of the NDP is to define profitable services for a set of demands, ports and a given fleet. The decisions of this problem are on a strategic level and do not handle how vessels should be sailed to these generated services. In the VSRP, the focus lies on recovering from disruptions due to weather conditions, port congestion or other reasons. For the recovery, an action is chosen that might increase costs and have impact on cargo and customers. The set of activities available to coordinators in the VSRP differs from the LSFRP, as does the timescale of the problem.

In contrast to liner shipping, Brønmo et al. (2007) and Korsvik et al. (2011) explain that tramp shipping does not require services and fixed schedules as tramp shipping companies engage in contracts for a specific amount of cargo. Therefore it is not possible to transfer problem-specific aspects from tramp shipping to the LSFRP even though the objectives of tramp shipping problems are similar.

## 4 Decision Support System

A DSS for the LSFRP should fulfill certain requirements in order to be relevant for the practical work of a repositioning coordinator. We propose a system using the simulated annealing approach in Tierney (2015) with some small modifications to allow for fast resolving. We list key requirements for the DSS in Table 1 from the perspective of the carrier, followed by a discussion of several visualization options, culminating with a recommendation for which one to use.

Data	Performance	Functions	Usability
Import/Export	Loading times	Planning	User Guidance
Storage	Fast optimization times	Optimization	Visualizations
Accessibility	System requirements	Reporting	Interactivity

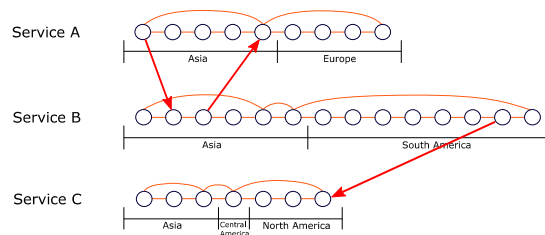
**Table 1: Requirements for a DSS for the fleet repositioning problem.**

One of the key challenges to deploying a DSS at a company is integrating it within the company's data management architecture. The LSFRP in particular draws on many different data sources within a company. We believe existing IT methodologies are sufficient for deploying our proposed DSS. Additionally, the DSS needs to obey certain system requirements and solve the LSFRP fast enough to allow the user to interact with the system.

The plans created by the DSS must be easy to use and modify. We propose several new visualizations of repositioning plans to realize this requirement. As Power and Kaparathi (2002) state, the development of the data model and the user interface for a web-based decision support system are the major tasks. Therefore we used the aforementioned aspects (ports, port sequences, services, networks, time, vessels, phase-out/phase-in and sail-on-services) for evaluating different visualizations of the fleet repositioning.

The spreadsheet view as presented by Fagerholt and Lindstad (2007) for the TurboRouter system, uses columns to represent time and vessels and individual cells to display port calls. With some changes, this view can be adapted to demonstrate services and their port sequences as well. Although this view displays time, it is not possible to include phase-outs, phase-ins and sail-on-services without making the view too complex. In addition to this, the connection between vessels and their services is not exactly clear, let alone the connections between services.

A simple approach to visualize services is to use a sequential graph for each service as is displayed in Figure 1. Each graph consists of nodes for the port calls and arcs for the sailings between calls. In order to add information about time, nodes and arcs can be labeled. Nevertheless, these labels lack the precision to determine the chronological order of phase-outs, phase-ins and the repositioning path.



**Figure 1: Sequential graphs to illustrate services and repositionings**

By connecting the separated graphs, it is possible to create a visualization that demonstrates the network structure of the services and the connecting ports (Figure 2). In this view, the repositioning coordinator also has an overview of possible ports for the repositioning path. Such a path could be displayed using additional arcs. However, in this type of view it is also problematic to add the dimension of time. Therefore the repositioning coordinator is not able to see the chronological relation between the possible port calls and the fixed schedules of the services.

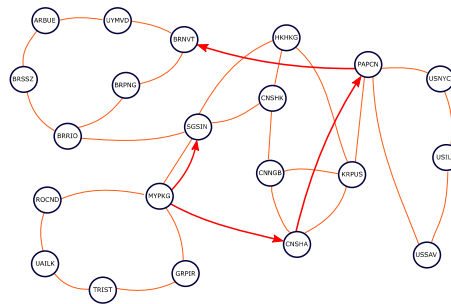


Figure 2: A connected graph to illustrate services and repositionings

A typical way to visualize networks of stations and paths is the “transit map” (Figure 3) as described by Guo (2011). This type of map is used for public transit plans to show the intersections of services and the relation between stations. Transit maps might give information about the time it takes to travel between two stations, but in the case of the LSRFP this information is not enough. In contrast to public transit, time distances between port calls are much longer. The repositioning coordinator needs precise information about the schedules in order to fully understand the repositioning plan, meaning this visualization provides insufficient information.

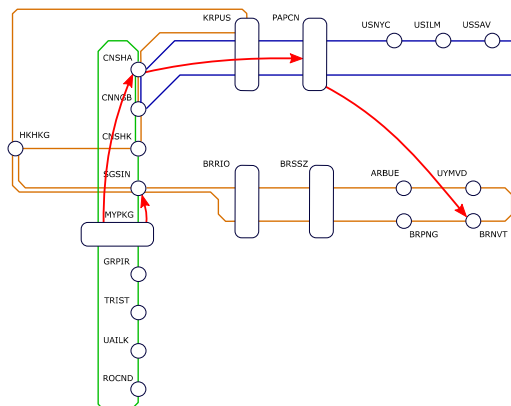
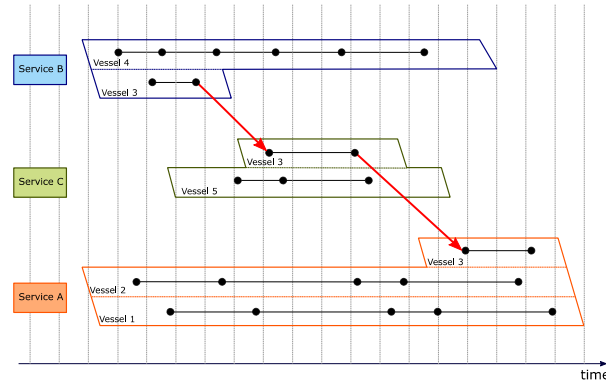


Figure 3: A transit map to illustrate services and repositionings

As the other approaches lack the possibility to include the time dimension, Figure 4 shows a visualization that is based on a time scale. By using this scale, the view presents the port calls of the services as nodes according to their occurrence. Additionally, arcs are used to represent sailings between ports. Services are displayed as closed boxes with multiple compartments, representing the associated vessels of each service. With these information, it is clear when the vessels are doing their port calls and how much time it takes to sail between two ports. There is no need for any kind of additional labels that might reduce the usability of the view. Repositionings can be visualized by

using extra arcs to show the port nodes for the phase-out and phase-in. Problem specific aspects like the sail-on-service opportunities can also be added to this view as can be seen in Figure 4.



**Figure 4: Our approach to illustrate services and repositionings**

Table 2 shows a summary of the previous comparison of the visualizations. The techniques of Figures 1-3 all lack the possibility to represent repositioning specific content like phase-outs (PO), phase-ins (PI) and sail-on-service opportunities (SOS). Only in Figure 4 is it possible to include these relevant information without making the view too complex or losing any usability.

	Ports	Sequence	Service	Network	Time	Vessels	PO/PI	SOS
Spreadsheet	✓	✓	✓	×	✓	✓	×	×
Fig. 1	✓	✓	✓	×	~	×	×	×
Fig. 2	✓	~	~	✓	×	×	×	×
Fig. 3	✓	✓	✓	✓	×	×	×	×
Fig. 4	✓	✓	✓	×	✓	✓	✓	✓

**Table 2: Comparison of the visualization techniques;**  
 ✓ can be displayed, x cannot be displayed, ~ can partly be displayed

## 5 Conclusion and future research

In this paper we examine different concepts to visualize relevant information of the liner shipping fleet repositioning problem in a domain-specific decision support system. Enhanced and intuitive visualizations will improve the general performance of coordinators, leading to economic benefits for liner shipping companies. We believe the visualizations we present are also relevant for other liner shipping problems. Furthermore, we seek to propose a web-based decision support system designed specifically for liner shipping fleet repositioning. The system should allow for an increased interactivity by enabling ad-hoc adjustments of solution processes.

## 6 Literature

Agarwal R, Ergun Ö (2008) Ship Scheduling and Network Design for Cargo Routing in Liner Shipping. *Transportation Science* 42(2):175–196

- Bausch DO, Brown GG, Ronen D (1998) Scheduling short-term marine transport of bulk products. *Maritime Policy & Management* 25(4):335–348
- Brønmo G, Christiansen M, Fagerholt K, Nygreen B (2007) A multi-start local search heuristic for ship scheduling—a computational study. *Computers & Operations Research* 34(3):900–917
- Brouer BD, Alvarez JF, Plum, Christian E. M., Pisinger D, Sigurd MM (2014) A Base Integer Programming Model and Benchmark Suite for Liner-Shipping Network Design. *Transportation Science* 48(2):281–312
- Brouer BD, Dirksen J, Pisinger D, Plum CE, Vaaben B (2013) The Vessel Schedule Recovery Problem (VSRP) – A MIP model for handling disruptions in liner shipping. *European Journal of Operational Research* 224(2):362–374
- Christiansen M, Fagerholt K, Nygreen B, Ronen D (2013) Ship routing and scheduling in the new millennium. *European Journal of Operational Research* 228(3):467–483
- Fagerholt K (2004) A computer-based decision support system for vessel fleet scheduling—experience and future research. *Decision Support Systems* 37(1):35–47
- Fagerholt K, Lindstad H (2007) TurboRouter: An Interactive Optimisation-Based Decision Support System for Ship Routing and Scheduling. *Marit Econ Logist* 9(3):214–233
- Fisher ML, Rosenwein MB (1989) An interactive optimization system for bulk-cargo ship scheduling. *Naval Research Logistics*(36):27–42
- Guo Z (2011) Mind the map! The impact of transit maps on path choice in public transit. *Transportation Research Part A: Policy and Practice* 45(7):625–639
- Korsvik JE, Fagerholt K, Laporte G (2011) A large neighbourhood search heuristic for ship routing and scheduling with split loads. *Computers & Operations Research* 38(2):474–483
- Power DJ, Kaparathi S (2002) Building Web-based Decision Support Systems. *Studies in Informatics and Control* 11(4):291–302
- Tierney K (2015) Optimizing Liner Shipping Fleet Repositioning Plans. *Operations Research/Computer Science Interfaces Series*, vol 57. Springer International Publishing, Cham
- Tierney K, Coles A, Coles A, Kroer C, Britt AM, Jensen RM (2012) Automated Planning for Liner Shipping Fleet Repositioning. *Proceedings of the 22<sup>nd</sup> International Conference on Automated Planning*:279-287
- Tierney K, Jensen RM (2012) The liner shipping fleet repositioning problem with cargo flows. *Lecture Notes in Computer Science Vol. 7555*:1-16
- UNCTAD (2014) Review of Maritime Transport 2014. United Nations Publications





# Mining Process Models from Natural Language Text: A State-of-the-Art Analysis

Maximilian Riefer<sup>1</sup>, Simon Felix Ternis<sup>2</sup>, and Tom Thaler<sup>2</sup>

<sup>1</sup> Saarland University, Institute for Information Systems (IWi), s9marief@stud.uni-saarland.de

<sup>2</sup> Institute for Information Systems at the German Research Center for Artificial Intelligence (DFKI GmbH) and Saarland University, {firstname.lastname}@iwi.dfki.de

## Abstract

Workflow projects are time-consuming processes. They include the knowledge extraction and the creation of process models. The necessary information is often available as textual resources. Therefore, process model mining from natural language text has been a research area of growing interest. This paper gives an overview of the current state-of-the-art in text-to-model mining. For this purpose, different approaches focusing on business process models are presented, analyzed and compared against each other on a theoretical and technical level. The resulting overview covers both advantages and disadvantages of current techniques. This should establish a sturdy basis on which further research can be conducted.

## 1 Introduction

Organizations are constantly trying to analyze and improve their business processes. This is only possible if the knowledge about the processes is available in a structured form like a business process model. The extraction of this knowledge and the generation of process models is a time-consuming and costly process. At the same time, 85% of the knowledge and information are estimated to be available in an unstructured form, mostly as text documents (Blumberg and Atré 2003). Modern natural language processing techniques render it possible to mine those text documents for process models. The automatic generation of process models from natural language text could speed up the whole workflow project of an organization. It gives people with no knowledge about process modeling the possibility to create process models, which is an important goal in view of the fact that structured data becomes more important. This field of research has been attracting more and more attention in recent years. Text Mining approaches have been developed for UML class diagrams (Bajwa and Choudhary 2011), entity relationship models (ERM) (Omar et al. 2008) or business process models (Friedrich et al. 2011). Current approaches do not aim at replacing an analyst but at helping him to create better models in less time.

The developed text mining approaches mainly focus on several specific model types and, thus, on specific contexts. A general overview in terms of a state-of-the-art analysis is missing. This paper

aims at giving an overview of existing approaches, the different natural language methods they use and a comparative analysis of the existing techniques. The focus of this paper is laid on the comparative analysis of the identified text mining approaches. Three main aspects are identified and analyzed: the textual input, the used NLP techniques and the model generation. Analyzing these aspects allows for a categorization of the existing approaches.

The used research methodology and the identification of the relevant literature are presented in section 2, while section 3 introduces the most important methods and terms for processing natural language texts. Section 4 gives an overview of current approaches. The comparative analysis, which consists of a detailed theoretical analysis and a proposed practical analysis, is conducted in section 5. The results are discussed in Section 6, followed by a conclusion in section 7.

## 2 Methodology

To define the current state of research and to identify different approaches for mining process models from natural language text, a systematical literature review was conducted. The three literature databases Google Scholar, SpringerLink and Scopus were used for the research. The following search keywords were derived from the title and thematic of this paper: *natural language processing*, *process model*, *process modeling* or *process model generation*, *process model discovery*, *text mining*, *process mining* and *workflow*. These were used in various combinations. As the used keywords cover broad research areas, they lead to a high number of search results. That complicated the identification of the relevant literature. The search results were checked through a title and abstract screening to identify the relevant work. Hence, only publications which explicitly mention text-to-model transformations were considered as relevant. There turned out to be a problem with the author's way of describing their work: instead of referring to text mining or natural language processing, they often used the text type, such as use cases or group stories, to outline their work. The search in a database was aborted when a significant amount of repetitions or loss of precision was noticed. It turned out, that the keyword search provided a low degree of relevant papers. Hence the keyword search was skipped and changed to a cross reference search. Table 1 shows the literature search results.

search results / database	number of results		
	Google Scholar	SpringerLink	Scopus
„process mining“ „natural language processing“	394	441	713
„process modeling“ „natural language“	5990	>10000	882
„process model“ „natural language“	>10000	>10000	2663
„process model discovery“ „natural language processing“	7	13	59
„process model generation“ „natural language processing“	50	21	169
„process mining“ „text mining“	525	575	1629
„workflow discovery“ „natural language processing“	18	3	7

**Table 1: Results of literature research**

Afterwards, further works were detected through a backwards search. The work of (Leopold 2013) provided a proficient starting point. The focus was set on approaches which generate a business process model. Five appropriate approaches could be identified:

1. BPMN model from text artefacts (Ghose et al. 2007)
2. BPMN model from group stories (Goncalves et al. 2009)

3. BPMN model from use cases (Sinha and Paradkar 2010)
4. BPMN model from text (Friedrich et al. 2011)
5. Model from text methodologies (Viorica Epure et al. 2015)

There seems to be a lack of publications from 2011 to 2015 and current works still refer to the aforementioned approaches (van der Aa et al. 2015). Because of the mentioned difficulties, it cannot be guaranteed, that the presented selection is exhaustive. The last approach does not build a BPMN model but is the most recent work deriving models from natural language texts which could be found. In the overview at hand, the different approaches will be presented and compared.

### 3 Introduction to Natural Language Processing

The processing of natural language requires different analyses of the text. This section introduces the most important methods and terms which are used for the identified approaches, whereby both syntactic and semantic analyses are performed. The syntactic analysis operates on a word at sentence level and annotates the words and phrases of a sentence, while the semantic analysis is used to provide an understanding of the semantic coherence of words and phrases in the text.

#### 3.1 Syntactic Analysis

The syntactic analysis usually comprises three main parts: the tokenization, the part-of-speech (POS) tagging and the parsing. *Tokenization* describes the process of splitting a text into several parts, called tokens. A token is a string between two separators like spaces, tabs or periods and, hence, usually consists of a single word. A *POS tagger* tries to annotate the different words with their word form (e.g. verb, noun). Many different methods for POS tagging are available. On average, they achieve an accuracy of about 97% of correctly annotated tokens (Manning 2011) in the context of natural language texts. Then, a *parser* analyzes a given sequence of tokens and builds a syntax tree, following a given grammar. It can be differentiated between a shallow parser and a deep parser. A shallow parser often just searches for nominal phrases and verbal phrases as smallest constituents, while a deep parser tries to determine every word and phrase of a sentence.

#### 3.2 Semantic Analysis

A semantic analysis is used to determine contextual meaning as e.g. homonyms or synonyms. For the semantic analysis, different semantic dictionaries and knowledge bases are used. Two well-known examples are WordNet (Miller 1995) and FrameNet (Baker and Sato 2003).

**WordNet.** “WordNet is a large lexical database of English. Nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms (synsets), each expressing a distinct concept. [...] The main relation among words in WordNet is synonymy, as between the words shut and close.”<sup>1</sup>

**FrameNet.** “The FrameNet project is building a lexical database of English that is both human- and machine-readable, based on annotating examples of how words are used in actual texts. [...]. FrameNet is based on a theory of meaning called Frame Semantic [...]. The basic idea is straightforward: that the meanings of most words can best be understood on the basis of a semantic frame: a description of a type of event, relation, or entity and the participants in it.”<sup>2</sup>

<sup>1</sup> Princeton University "About WordNet." WordNet. Princeton University. 2010. <<http://wordnet.princeton.edu>>

<sup>2</sup> <https://framenet.icsi.berkeley.edu/fndrupal/about>

Another part of semantic analysis is the anaphora resolution. An anaphora “is a word that refers to a word used earlier in a sentence and replaces it, for example the word 'it' in the sentence 'Joe dropped a glass and it broke’<sup>3</sup>. Anaphora resolutions links noun phrases to the related pronoun. This is also important for text-to-model mining, because activities can be split up into different sentences and anaphora analysis helps to connect and recognize these sentences as one activity.

## 4 Overview on Existing Approaches

This section introduces the different approaches. Every approach is categorized in three aspects. Textual input refers to the required source document and its constraints. Text analysis introduces the used NLP technologies. Model generation describes the resulting model characteristics.

### 4.1 Process Discovery from Model and Text Artefacts (Ghose et al. 2007)

**Textual input.** As there are no explicit restrictions mentioned for text files, every textual input should be usable. Additionally, this approach is able to use existing BPMN models to check for consistency of the created models.

**Text analysis.** A text is analyzed in two ways. One is to look for text patterns which indicate process structures. Such a pattern could be “If < condition/event >, [then] < action >“. A second analysis is performed with the Natural Language Toolkit (NLTK) (Bird et al. 2009). First, the text is annotated with POS tags and, then, a syntax tree is parsed by a shallow parser. In this tree, combinations of verb phrases (vp) and noun phrases (np) indicate activities (a) and actions. The sequence <vp,np> would be considered as an activity where (np) functioned as an object in difference to an action sequence <np, a> where (np) functions as an actor. Sequence flows are discovered through predefined signal words.

**Model generation.** No sound BPMN model is created. Discovered model parts are modeled without connection if no coherence is found.

### 4.2 Business process mining from group stories (Goncalves et al. 2009)

**Textual input.** The authors use group stories, a collective narrative technique, to extract knowledge from people who are involved in a process (Santoro et al. 2008). A story which is written by a team is likely to contain more useful information than a collection of individual interviews. A team consists of story tellers, moderators and analysts. A moderator is responsible for a correct and usable text format. The required form of input texts is not specified any further.

**Text analysis.** First, the text is tokenized, then, it is annotated with POS tags by a trigram tagger (Brants 2000) and parsed by a shallow parser. Activities, actors, actions and parameters are identified through templates based on patterns of verb- and noun-phrases. Flow elements are identified through keywords. No semantic analysis is performed. Due to the similarity of process models to scenarios, they use a variation of the CREWS scenario model to store the extracted information. This is based on a CREWS use case, which showed similarities to process descriptions. Both actions and actors are core concepts of a model. An action is performed by an actor and can have one or more start and end events (Maiden 1998).

---

<sup>3</sup> <http://dictionary.cambridge.org/dictionary/english/anaphor>

**Model generation.** The extracted model elements are modeled into BPMN models. These models do not have to be complete, e.g. they don't necessarily have start or end events. With the help of analysts, the story telling team generates the final model.

#### 4.3 Use Cases to Process Specifications in Business Process Modeling Notation (Sinha and Paradkar 2010)

**Textual input.** Use cases are used as source documents. The authors did not provide a required form for use cases, the process descriptions have to be extracted manually. They define a use case as a sequence of statements, statements as a sequence of actions. A statement can be a conditional statement, which is an exception statement associated to the nearest preceding statements.

**Text analysis.** The text is tokenized and lemmatized. Afterwards, the POS tagging is conducted by a tagger based on a linear classification technique called robust risk minimization (Zhang et al. 2002). The text is parsed by a shallow parser based on "Finite State Transducer" (Boguraev and Neff 2008). For the semantic analysis with an anaphora resolution, domain specific databases are created manually. Flow elements are identified by keywords. The gained information is stored in a use case description meta model containing a model of the domain with its actors and business items and a use case model, in which the actions and initiating actors of every sentence are stored.

**Model generation.** For every actor, a swim lane is created. Afterwards, the process elements are added and the process model is built sentence by sentence. The approach only supports a basic selection of BPMN symbols. Furthermore, a method is presented, which allows for an automated combination of single process models resulting in a complete business process.

#### 4.4 Process Model Generation from Natural Language Text (Friedrich et al. 2011)

**Textual input.** Every text can be used as long as it fulfills some requirements. The text has to describe a model, must not contain questions and has to be described sequentially. Furthermore, it should describe the view of an involved actor and non-sequential jumps have to be made explicit.

**Text analysis.** For the syntactical analysis, tools from the Stanford NLP Group are used to parse a syntax tree with Stanford dependencies. The semantic analysis is based on dictionaries and databases as WordNet and FrameNet and also includes a self-developed anaphora resolution. The authors utilize stop word lists to identify flow elements. As in earlier approaches, they use a variation of the CREWS scenario model as a world model to store the intermediate information.

**Model generation.** After the creation of swim lanes for the actors, nodes for every activity and event are created. These nodes are connected with their related flow elements. This approach is able to use more BPMN elements thanks to a deeper and more complex text analysis.

#### 4.5 Automatic Process Model Discovery from Textual Methodologies (Viorica Epure et al. 2015)

**Textual input.** This approach uses archaeological methodologies as textual inputs. Such methodologies describe archaeological processes with natural language texts. The authors only chose methodologies which describe a single process instance to simplify the given task.

**Text analysis.** In a first step, the text is cleaned and unwanted punctuation and phrases are removed. The Stanford Parser and a combination of Stanford and NLTK tagger are used to generate a syntax tree. Afterwards, transitive verbs are identified as they are the elements most likely to represent an activity. For this task, the databases WordNet and VerbNet (Kipper et al. 2004) are used. To identify

the relationships between activities, domain specific rules, based on patterns of found activities and keywords, are utilized. The text is analyzed sentence by sentence.

**Model generation.** The only identified elements are activities and relationships, no explicit model is generated. A textual representation is returned (e.g. Start → (Act1 || act2) → Act4 → Stop).

## 5 Comparative Analysis

The comparative analysis has to be performed on a theoretical basis because there are no implementations available. The theoretical analysis bases on the previous presentation of the approaches. The different aspects *textual input*, *text analysis* and *target document* are compared and, if possible and useful, they will be rated. For lack of implementations, the involved research groups mostly use internal prototypes or refer to software and applications which were no longer accessible or operable. Hence, a practical analysis was not feasible. The insights obtained by the overview of existing approaches will be used to propose some measures and automated techniques to analyze and evaluate future approaches with the corresponding implementations.

### 5.1 Theoretical Analysis

There are three main aspects which can be analyzed and compared. These are the textual input, the chosen NLP technique and the procedure of deriving a process model. In the following analysis, the different approaches, in order of the given overview, are referenced as Artefacts, GroupStories, UseCases, Alltext and Methodologies.

#### 5.1.1 Textual Input

Artefacts is the only approach which does not introduce any constraints for input texts. This is possible because it is not the goal to extract a complete model but just model parts in no particular order to be combined by an analyst. GroupStories utilizes a moderator during the extraction of knowledge to guarantee a reliable process description. In UseCases, it is not explicated which form a use case is required to have. They state that the relevant parts are extracted manually with all numeral extensions but not whether the order of the sentences is altered. It is to assume that it, like all other approaches, depends on a sequential order of the event descriptions and an explicit indication of jumps. Overall, the text requirements introduced for Alltext generally apply to every approach which ties to generate a correct model out of a model description. These requirements are, besides the sequential order, the non-presence of questions in the process description and the existence of a process level.

#### 5.1.2 NLP techniques

The core of every text-to-model approach is the text analysis with NLP. The amount and quality of extracted information decides on the complexity of the generated models. Besides the ability to extract the information of a given problem, an important factor of NLP techniques is their flexibility. To prove useful as a universal text-to-model approach, the capability to adapt to different domains or even languages is necessary.

As a first step, every approach conducts a syntactic analysis. They use POS taggers to annotate the text. Alltext and Methodologies use the Stanford Tagger, Artefacts the NLTK, GroupStories a trigram tagger and UseCases a tagger on basis of Robust Risk Minimization (RRM). POS tagging

is a well-researched problem. With exception of RRM (92%), all taggers work with a precision of approx. 97%. Methodologies found that a combination of tagging methods could yield to a better result. Therefore, the chosen technique is of minor importance. The POS taggers are available for many languages or can easily be trained for different languages with respective text corpora.

In a second step, the annotated texts are parsed to build a syntax tree. Artefakt, GroupStories and UseCases use shallow parsers. Alltext and Methodologies use, with the Stanford Parser, a deep parsing method. With more complex trees, better results can be achieved. Except for GroupStories, all approaches execute a semantic analysis. Alltext and Methodologies use available databases and dictionaries as WordNet, FrameNet or VerbNet. Artefacts and UseCases rely on self-created domain databases. By creating own databases, they achieve a high flexibility and can easily be adapted to different domains. UseCases and Alltext implemented concepts of anaphora resolution as part of the semantic analysis. This allows the detection of activities which are described in more than one sentence.

With all information available and all elements identified, the produced syntax trees were searched for patterns of verb and noun phrases to identify activities, events and actors. The activities are the central elements of a process model. To identify these activities, verbs and verb phrase and their related noun phrases are used. For the identification of events and sequence flows, word lists which indicate sequential (e.g. then, after), conditional (e.g. if, whereas) or parallel (e.g. while, in the meantime) flows were established. Methodologies uses a domain specific database with all combinations of verbs and signal words to identify model elements. Moreover, GroupStories, UseCases and AllText use a modified meta model based on the CREWS scenario model to store the extracted information.

### 5.1.3 Model Generation

Methodologies only analyses the text for activities and does not create a BPMN model but a descriptive process representation out of activities (e.g. Start → (Act1 || act2) → Act4 → Stop). Artefacts only creates model parts, which have to be combined by an analyst. For the remaining approaches, the extracted data, stored in a modified CREWS meta-model, is used to build the process models. First, an actor will be represented as swim lane in a BPMN model. Then, events and activities will be included. Alltext creates all nodes first and connects them with sequence flows afterwards. UseCases creates the activities, events and sequence flows in the order of their occurrence in a use case. Because of its advanced text analysis, Alltext is able to use a higher number of BPMN symbols and therefore can create a more complex process model.

AllText is the only approach which refers to current labeling conventions and applies them to label the model elements appropriately. Methodologies mentions the labeling problem but does not stick to the conventions.

### 5.1.4 Summary

The approaches are different in most of their aspects. Thus, they provide a good overview of current methods and possibilities of text-to-model mining. Table 2 summarizes the previously presented sections. The resulting advantages and disadvantages of the presented approaches are listed in Table 3. The most advanced among the presented approaches is (Friedrich et al. 2011). It introduces many features and ideas, which should serve as standard for every approach.

			Approach				
			Artifacts (Ghose et al. 2007)	Group-Stories (Goncalves et al. 2009)	UseCases (Sinha and Paradkar 2010)	Alltext (Friedrich et al. 2011)	Methodology (Viorica Epure et al. 2015)
Textual Input	Flexi- bility	Speech		•	•		
		Domain			•	•	
	Require- ments	Sequence			•	•	•
		Explicit Jumps				•	
		Keywords	•	•	•	•	
NLP-Techniques	Syntactic Analysis	Stanford-Tagger				•	•
		NLTK-Tagger	•				•
		Trigram-Tagger		•			
		Linear Classific.			•		
		Stanford-Parser				•	•
		NLTK-Parser					
		Shallow- Parser	•	•	•		
		Templates	•				
	Semantic Analysis	WordNet				•	•
		FrameNet				•	
		VerbNet					•
		Individ. Ontologies	•		•		
		Anaphora Res.			•	•	
Model Generation	Structured Desc.			•	•		
	Activities	•	•	•	•	•	
	Connectors / Gateways	•	•	•	•	•	
	Events			•	•		
	Actors	•	•	•	•		

Table 2: Comparative analysis summary

Although some ideas can also be found in earlier approaches, Friedrich combined them to create the most complete approach. As for texts, it introduces only minor constraints to support the text mining. For the text analysis, state-of-the-art applications as the Stanford Parser and databases as WordNet were used. Additionally, an own technique for anaphora resolution was presented, which, overall, achieved a better score than existing anaphora resolution applications as BART (Versley et al. 2008) in this special environment of process models. Due to the thoroughly executed analysis, a more complex model generation is rendered possible. On average, they were able to correctly reconstruct 76% of the models in a test set based on their textual descriptions. Meanwhile, the used NLP techniques are outdated. Usually, an approach should advance with the possibilities of its NLP techniques. Hence, a new evaluation with current techniques is necessary.

	Advantages	Disadvantages
<b>Artefacts</b>	use of reference models	creates only model parts
<b>GroupStories</b>	flexible about speech	no semantic analysis no complete BPMN models
<b>UseCases</b>	very flexible about speech and domain allows relative fast computation	no clarity about input text format high effort and domain knowledge necessary
<b>AllText</b>	creates complete models thorough semantic analysis	-
<b>Methodologies</b>	no word lists necessary	creates no models

Table 3: Advantages and disadvantages of the available approaches



## 5.2 Proposed Practical Empirical Analysis Approaches

The different approaches presented try to build a model from natural language text, hence, the quality of the generated model can be used to rate the performance of an approach. However, it cannot be stated how well an approach performs on a given text. To compare the general performance, a test set has to be used. Such a test set contains text-model pairs. These pairs are validated by experts and analysts. The automatically generated model and the reference model can be compared to measure the performance. A well-designed approach should produce a high quality process model. With further improvements of the developed approaches and after reaching a certain standard for the generated models, the run time of an approach should be as low as possible. Therefore, the time needed to create a model out of a textual resource should be part of every practical analysis. There is a lack of available implementations for the presented approaches and a practical analysis is not feasible. Such an analysis should be the content of future research. For this reason and encouraged by the evaluation of Friedrich et al. (2011), a research group at the DFKI is currently working on an advanced version of this approach. As a part of the RefMod-Miner<sup>4</sup>, a text-to-BPMN and a newly adapted text-to-EPC approach are realized. Currently, the applications are in a testing stage. They show some promising results when the textual resources fit certain textual requirements. Overall, there is a low tolerance regarding different text types. In the future, a practical analysis based on the proposed techniques will be performed.

### 5.2.1 Model Quality

Recent research states the importance of model quality. There are two rather easy to check quality criteria: The syntactic quality, which states whether the model is correctly modeled according to the syntax of the modeling language and the semantic quality, which states how well a model describes the modeled domain. Since the models are generated automatically, the syntactic quality should be guaranteed. The semantic model quality of an approach can be measured with the help of a test set with texts and their corresponding manually generated models. In combination with the model similarity, a metric can be computed serving as a measure for the semantic quality. A third, rather subjective, criterion is the practical quality, which defines whether a model can easily be understood. (Houy et al. 2012) did some research on how to measure the understandability of conceptual models like process models. They found that understanding a model contains different dimensions. The aspects of understanding a model can be subjective or objective. Furthermore, the understanding can depend on the effectiveness or the efficiency of the model. Four more or less measurable understanding criteria were identified: *correctly answering questions about the model*, *problem solving based on the model content*, *perceived ease of understanding a model* and *time needed to understand a model*. These criteria cannot be computed automatically, which makes it hard to rate an approach based on how understandable the created model is. But it has been shown that the understanding of process models can be improved by a systematic labeling style. Therefore, an approach should at last follow naming conventions. (Leopold et al. 2013) introduced an approach to automatically detect naming convention violations. Moreover, the model quality has to be seen in reference to the participants of a process and the creators of a process description. Thus, terminology, grammatical structures or abstraction levels can differ. For example, the description and perception of a process by the head of the sales department may significantly differ from the technical specification used by a system administrator.

---

<sup>4</sup> <http://refmod-miner.dfki.de>

### 5.2.2 Model Similarity

If reference models are available, the automatically and manually generated models can be compared to determine the performance of an approach. A reasonable metric for this task is the model similarity, which can be applied for different aspects of a process model (Dijkman et al. 2011), as e.g. based on structural properties, on semantic similarities of the labels or on both aspects. In the area of business process models, there exist various similarity measures which can be used for the practical empirical evaluation. The selection of the similarity measure has to be carried out depending on the content of the automatically generated model. If the model contains, for example, organizational units, the similarity measure should consider these for the similarity calculation. Otherwise, it is not possible to correctly verify the quality of the derived model. The reference model can be seen as an optimal description in form of a process model of the input text. Thus, the calculation of the similarity value between the manually generated reference model and the automatically generated process model provides a numerical value, which can be used to determine the overall quality of the generated model. A higher similarity value means a higher conformance and, thus, a higher quality of the generated model.

## 6 Discussion

With the presentation and the comparative analysis of current approaches, a starting point for further research was given. The previous analysis was conducted on a theoretical basis as there are not enough implementations available. The research groups are using internal prototypes. The analysis could be much more informative if a proper practical analysis could be executed. Furthermore, with the focus on business process models, only a small area of the topic is covered. Still, it shows the difficulties of further research. Differences in possible textual resources and the different forms of output models require many different approaches. An important factor in future research is the importance of the individual research parts, which have to be combined for text-to-model transformations. Mainly, those are NLP, process models and model mining. Further research and development will benefit from the work at hand and shall guarantee advancement.

A simple analysis was performed by the authors, with an independent text-to-EPC implementation, based on Friedrich et al. (2011), which is considered as the current state-of-the-art. It was not validated with a full test set but small tests already showed some existing problems. The low tolerance towards flawed textual input files is a problem for current and future research. Small differences in the required text format results in a significantly lower model quality, not taking spelling and punctuation errors into account. That could be a reason for which the research on text-to-model transformation was declining. The requirements for textual process descriptions are too high. Further research is required to handle more variations and errors of textual inputs.

## 7 Conclusion

The goal of this paper was to determine the state-of-the-art of the current research in text-to-model transformation. The major focus was set on approaches which generate business process models. Since 2011, no thorough approach has been published and only one more current but incomplete approach from 2015 could be found. A brief presentation of existing approaches was given and combined with a theoretical comparative analysis. Due to a lack of available implementations, no practical analysis could be performed. Although the lack of a practical analysis, the approach

developed by Friedrich was identified as the current state-of-the-art. Thus, the current research project at the DFKI aims at an advanced implementation of (Friedrich et al. 2011). As part of the research prototype RefMod-Miner, BPMN models and EPCs can be generated. The application shows some of the problems which have to be overcome in the next years. This paper presents a first step towards further research in the field of model mining out of natural language text. In the future, the focus should be set on an increased text tolerance in order to enable the analysis of more complex textual resources and to gain a deeper understanding of them. Another goal should be to obtain more implementations and focus on a more practical analysis of existing approaches. In this way, strengths and weaknesses of approaches could be identified more reliably and the gained information used to develop more advanced methods.

## 8 Literature

- Bajwa IS, Choudhary MA (2011) From Natural Language Software Specifications to UML Class Models. Paper presented at the 13th International Conference on Enterprise Information Systems
- Baker CF, Sato H (2003) The FrameNet data and software. Paper presented at the Proceedings of the 41st Annual Meeting on Association for Computational Linguistics - Volume 2, Sapporo, Japan
- Bird S, Klein E, Loper E (2009) Natural Language Processing with Python. O'Reilly Media, Inc.,
- Blumberg R, Atre S (2003) The problem with unstructured data. *RM Review* (13):42-49
- Boguraev B, Neff M (2008) Navigating through Dense Annotation Spaces. Paper presented at the Sixth International Conference on Language Resources and Evaluation (LREC'08), Marrakech, Morocco
- Brants T (2000) TnT: a statistical part-of-speech tagger. Paper presented at the Proceedings of the sixth conference on Applied natural language processing, Seattle, Washington,
- Dijkman R, Dumas M, van Dongen B, Käärik R, Mendling J (2011) Similarity of business process models: Metrics and evaluation. *Information Systems* 36 (2):498-516. doi:Doi 10.1016/J.Is.2010.09.006
- Friedrich F, Mendling J, Puhmann F (2011) Process model generation from natural language text. Paper presented at the Proceedings of the 23rd international conference on Advanced information systems engineering, London, UK
- Ghose A, Koliadis G, Chueng A Process Discovery from Model and Text Artefacts. In: *Services, 2007 IEEE Congress on*, 9-13 July 2007 2007. pp 167-174. doi:10.1109/services.2007.52
- Goncalves JC, Santoro FM, Baiao FA (2009) Business process mining from group stories. Paper presented at the Proceedings of the 2009 13th International Conference on Computer Supported Cooperative Work in Design, CSCWD, Santiago
- Houy C, Fettke P, Loos P (2012) Understanding Understandability of Conceptual Models – What Are We Actually Talking about? In: Atzeni P, Cheung D, Ram S (eds) *Conceptual Modeling, 31st International Conference ER 2012 (LNCS 7532)*. Lecture Notes on Computer Science, vol 7532. Springer, Berlin, Heidelberg, pp 64-77
- Kipper K, Snyder B, Palmer M (2004) Extending a Verb-lexicon Using a Semantically Annotated Corpus. Paper presented at the LREC

- Leopold H (2013) Natural Language in Business Process Models. Lecture Notes in Business Information Processing (LNBIP), vol 168. Springer, Berlin
- Leopold H, Eid-Sabbagh R-H, Mendling J, Guerreiro Azevedo L, Araujo Baião F (2013) Detection of naming convention violations in process models for different languages. *Decision Support Systems* 56 (1):310-325
- Maiden NAM (1998) CREWS-SAVRE: Scenarios for Acquiring and Validating Requirements. In: Sutcliffe A, Benyon D (eds) *Domain Modelling for Interactive Systems Design*. Springer US, pp 39-66. doi:10.1007/978-1-4615-5613-8\_3
- Manning CD (2011) Part-of-speech tagging from 97% to 100%: is it time for some linguistics? Paper presented at the Proceedings of the 12th international conference on Computational linguistics and intelligent text processing - Volume Part I, Tokyo, Japan
- Miller GA (1995) WordNet: a lexical database for English. *Communications of the ACM* 38 (11):39-41
- Omar N, Hassan R, Arshad H, Sahran S (2008) Automation of database design through semantic analysis. Paper presented at the Proceedings of the 7th WSEAS international conference on Computational intelligence, man-machine systems and cybernetics, Cairo, Egypt,
- Santoro FM, Borges M, Pino JA Tell us your process: A group storytelling approach to cooperative process modeling. In: *Computer Supported Cooperative Work in Design, 2008. CSCWD 2008. 12th International Conference on, 16-18 April 2008 2008*. pp 29-34. doi:10.1109/cscwd.2008.4536950
- Sinha A, Paradkar A (2010) Use Cases to Process Specifications in Business Process Modeling Notation. Paper presented at the 2010 IEEE International Conference on Web Services, Miami, Florida
- van der Aa H, Leopold H, Mannhardt F, Reijers H (2015) On the Fragmentation of Process Information: Challenges, Solutions, and Outlook. In: Gaaloul K, Schmidt R, Nurcan S, Guerreiro S, Ma Q (eds) *Enterprise, Business-Process and Information Systems Modeling*, vol 214. *Lecture Notes in Business Information Processing*. Springer International Publishing, pp 3-18. doi:10.1007/978-3-319-19237-6\_1
- Versley Y, Ponzetto SP, Poesio M, Eidelman V, Jern A, Smith J, Yang X, Moschitti A (2008) BART: a modular toolkit for coreference resolution. Paper presented at the Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies: Demo Session, Columbus, Ohio
- Viorica Epure E, Martin-Rodilla P, Hug C, Deneckere R, Salinesi C Automatic process model discovery from textual methodologies. In: *Research Challenges in Information Science (RCIS), 2015 IEEE 9th International Conference on, 13-15 May 2015 2015*. pp 19-30. doi:10.1109/rcis.2015.7128860
- Zhang T, Damerau F, Johnson D (2002) Text chunking based on a generalization of winnow. *J Mach Learn Res* 2:615-637

# Konstruktion eines adaptiven Referenzmodells für die Materialflussgestaltung in der Produktionslogistik

Alexander Schubel<sup>1</sup>, Christian Seel<sup>2</sup> und Markus Schneider<sup>1</sup>

<sup>1</sup> Hochschule Landshut, Kompetenzzentrum PuLL,  
{alexander.schubel | markus.schneider}@haw-landshut.de

<sup>2</sup> Hochschule Landshut, IPIM, christian.seel@haw-landshut.de

## Abstract

Im Rahmen ganzheitlicher Produktionssysteme gelten die Prinzipien und Methoden der Lean Production als Maßstäbe für die effiziente Gestaltung der Prozesse. In Bezug auf die aktuellen Herausforderungen bei der Gestaltung einer Lean Production bietet die Referenzmodellierung einige Potentiale für die Produktions- und Logistikprozessplanung. Der vorliegende Beitrag legt die Notwendigkeit eines geeigneten Referenzmodelles dar und beschreibt die Materialflussgestaltung der Produktionslogistik in Form eines adaptiven Referenzmodells. Abschließend werden die Adaptionenmechanismen des Referenzmodells durch einen Software-Prototypen veranschaulicht.

## 1 Herausforderungen in der Produktions- und Logistikplanung

Die einem Produktionssystem zugrunde liegenden Prinzipien haben seit Anfang des zwanzigsten Jahrhunderts einen rasanten Wandel erfahren (Westkämper 2013). Für die Wirtschaftsinformatik sind dadurch sowohl neue Herausforderungen in der Organisationsgestaltung in Bezug auf Aufbau- und Ablauforganisation sowie deren IT-Unterstützung entstanden. Als aktueller Stand der Technik für die variantenreiche Serienproduktion können die Prinzipien und Methoden der Lean Production und damit als Maßstäbe für die effiziente Gestaltung moderner Produktionssysteme angesehen werden (Günthner und Boppert 2013). Im Gegensatz zu anderen Domänen, wie dem Handel, für die bereits verbreitete Referenzmodelle, wie das Handels-H (Becker und Schütte 2004) existieren, fehlt für Lean Production ein solches Referenzmodell, das zur Organisationsgestaltung sowie deren IT-Unterstützung verwendet werden kann. Konkreter lassen sich in den Ausarbeitungen zum Thema Lean Production aktuell drei Defizite feststellen. Es fehlt an einer *konsistenten, vernetzten und aufeinander bezogen Darstellung* des Themas, an einer *ausreichenden Formalisierung in Form von Informationsmodellen* sowie an einer *unternehmensübergreifenden neutralen Darstellung*, die auch von kleinen und mittleren Unternehmen (KMUs) genutzt werden kann (Schubel et al. 2015b).

## 2 Potentiale der Referenzmodellierung in der Produktions- und Logistikplanung

Aufgrund dieser drei Defizite bietet sich das Gebiet der Informationsmodellierung, genauer gesagt der Referenzinformationsmodellierung vielversprechende Lösungsansätze.

Ein Referenzmodell für Lean Production ermöglicht durch eine geeignete semi-formale und standardisierte Darstellung ein eindeutigeres und intersubjektiv nachvollziehbares Verständnis der Thematik in Wissenschaft und Praxis. Somit wird durch die Verwendung semi-formaler Informationsmodellierungssprachen der Formalisierungsgrad der häufig nur textuell vorliegenden Lean Production-Ansätze gesteigert und auch in Modellierungswerkzeugen erfassbar gemacht.

Ferner ergeben sich Vorteile im Wissensmanagement sozio-technischer Systeme durch Formalisierung. Diese unterstützt die Wissensverteilung, -nutzung und -bewahrung (Staiger 2008).

Zudem sind die Potentiale und der Nutzen der Referenzmodellierung im Bereich der Produktions- und Logistikplanung sehr hoch einzuschätzen, da einerseits die Effektivität und Effizienz von praktischen Planungsvorhaben gesteigert werden können (Schmelzer und Sesselmann 2013; Thomas und Scheer 2006). Andererseits wird die Lean-Philosophie als Ingenieursdisziplin weiterentwickelt, indem die Potentiale der Wiederverwendbarkeit anhand von Referenzmodellen erhöht werden (Fettke und Loos 2005).

Im Kontext der Industrie 4.0 schafft die rechnerkonforme Syntax, durch eine semi-formale Darstellung, die Grundlagen für die weitere Digitalisierung und Automatisierung der Prozessplanung des physischen Materialflusses. Die im Zuge der Referenzmodellierung erreichte Prozessstandardisierung erhöht zusätzlich die Wandlungsfähigkeit der modellierten Produktionssysteme (Krebs et al. 2011). Eine wandlungsfähige Produktion und Logistik gelten ebenfalls als Anwendungsfelder der Industrie 4.0 (Bauernhansl et al. 2014). Des Weiteren können *adaptive* Referenzmodelle (Delfmann 2006) ein konkretes Lösungskonzept für die geforderte flexible Prozessstandardisierung im Rahmen von modernen Produktionssystemen darstellen (Binner 2010).

## 3 Forschungsziel und -methodik

Aus den aufgezeigten Potentialen der Referenzmodellierung in der Produktions- und Logistikplanung ergibt sich die *Entwicklung eines adaptiven Referenzmodells für die Produktionslogistikgestaltung bei KMUs* als Forschungsziel.

Es lassen sich in der Wirtschaftsinformatik zwei Gruppen von Forschungsfragen unterscheiden. Die erste Gruppe umfasst Forschungsfragen, welche darauf abzielen, den gegenwärtigen Einsatz und die Entwicklung von Informationssystemen sowie die dazu angewendeten Methoden zu untersuchen. Die zweite Gruppe bilden Forschungsfragen, welche die Entwicklung neuer Lösungen, im Sinne einer „Design Science“, zum Ziel haben (Hevner und Chatterjee 2010). Im Vorfeld der Entwicklungsarbeiten gilt es, die Forschungslücke durch behavioristische Forschungsansätze zu konkretisieren. Als Ergebnis einer Literaturanalyse sind dafür die Entwicklungspotentiale am aktuellen Referenzmodellbestand dargestellt, welche anschließend durch eine empirische Anforderungserhebung im Speziellen bestätigt werden (Kapitel 4). Die darauffolgende Entwicklung des adaptiven Referenzmodells orientiert sich methodisch an den Konstruktionsschritten nach (Fettke und Loos 2004):



**Bild 1: Entwicklungsmethodik, in Anlehnung an (Fettke und Loos 2004)**

Die *Problemdefinition* erfolgt im Rahmen der Einleitung (Kapitel 1) und Konkretisierung der Forschungslücke (Kapitel 4) sowie bei der Auswahl der Modellierungssprache in Kapitel 5. Die *Konstruktion im engeren Sinne* wird ebenfalls in Kapitel 5 beschrieben. Das Referenzmodell gilt es, laufend zu *bewerten* (formative Evaluation). Somit fließen die Evaluierungsergebnisse wieder direkt in die weiteren Konstruktionsarbeiten ein. Dies erfolgt durch die Erprobung der Prozesse des adaptiven Referenzmodells in einem mittelständischen Unternehmen. Dabei werden sowohl wirtschaftliche als auch technische Aspekte berücksichtigt. Außerdem sind Anforderungen aus der empirischen Erhebung für die Evaluierung heranzuziehen. Die *Pflege* des Referenzmodells gilt es einerseits durch anschließende Forschungsvorhaben sicherzustellen, andererseits ist in der Praxis die kontinuierliche Referenzmodellpflege durch beispielsweise den Einsatz benutzerfreundlicher Modellierungswerkzeuge zu begünstigen.

#### 4 Entwicklungspotentiale und Anforderungen im Planungsumfeld der Produktionslogistik von KMUs

Der anhand einer umfangreichen Literaturanalyse identifizierte Referenzmodellbestand in der Produktions- und Logistikplanung weist die folgenden Defizite als Ausgangspunkt für die Referenzmodellentwicklung auf (Schubel et al. 2015b):

- Keine Referenzprozessmodelle zur Ausgestaltung des physischen Materialflusses moderner Produktionssysteme.
- Konzepte und Lösungsmöglichkeiten zur Prozessgestaltung mit einem niedrigen Formalisierungsgrad und ohne Adaptionsmechanismus.
- Fehlende konsistente, vernetzte und aufeinander bezogene Darstellung der einzelnen Sachverhalte der Lean Production.
- Branchenübergreifende bzw. -neutrale Darstellung fehlt (Schwerpunkt der Anwendungsfälle: Automobilhersteller).

Um die praktische Relevanz der Potentiale zu bestätigen erfolgte eine empirische Anforderungserhebung im Rahmen von acht Experteninterviews (Schubel et al. 2015c). In Bezug auf die Defizite des Referenzmodellbestandes umfasst die Grundgesamtheit der befragten Experten Prozessplaner in KMUs. Im Speziellen sind Planungsverantwortliche der Produktionslogistik geeignet, da im Rahmen der Produktionslogistik der physische Materialfluss und die zugehörigen Informationsflüsse vom Wareneingang über den gesamten Produktionsprozess hinweg bis zum Warenausgang zu planen sind (Pawellek 2007). Somit stellt die Produktionslogistik eine wesentliche Querschnittsfunktion „von Rampe zu Rampe“ eines produzierenden Unternehmens dar. Zusammenfassend hat die Produktionslogistik maßgeblichen Einfluss auf die Gestaltung eines schlanken und modernen Produktionssystems.

#	Anforderungen	Anzahl eingeflossener Interviewstimmen
A.1	Systematische Prozessplanung	86
A.2	Verfügbarkeit von wiederverwendbaren und formalisierten Referenzlösungen für die Prozessplanung	80
A.3	Reduzierung von Komplexität und Verschwendung durch die Prozessplanung	69
A.4	Identifizierbarkeit von Einsparungspotenzialen	57
A.5	Transparenz der Gesamtprozesse	50
A.6	Formalisierung von implizitem und explizitem MA-Wissen	36
A.7	Anwenderfreundliche Methoden und Modelle	28
A.8	Funktionsübergreifende Prozessplanung	24
A.9	Wertschöpfungsorientierte Prozessplanung	8
A.10	Eindeutige Organisationsgestaltung	6
A.11	Konfigurationshilfen für die Prozessplanung	4
A.12	Anwendungsflexible Methoden und Modelle	3
A.13	Unterstützung durch IT-Systeme	2

**Tabelle 1: Anforderungen in der Produktionslogistik von KMUs (Schubel et al. 2015c)**

Tabelle 1 stellt die für die Referenzmodellierung relevanten Anforderungen dar. Für den Anwendungsfall der Produktionslogistikgestaltung bei KMUs können die praktischen Anforderungen als bekräftigend für die Relevanz der Entwicklungspotentiale der Referenzmodellentwicklung bewertet werden (Schubel et al. 2015c).

## 5 Entwicklungsstand des adaptiven Referenzmodells für die Produktionslogistikgestaltung

In Bezug auf die Potentiale und spezifischen Anforderungen erfolgt die Konstruktion eines *adaptiven und semi-formalen Referenzmodells für die Prozessplanung des physischen Materialflusses in der Produktionslogistik von KMUs*. Die folgenden Ausführungen beschreiben den aktuellen Entwicklungsstand.

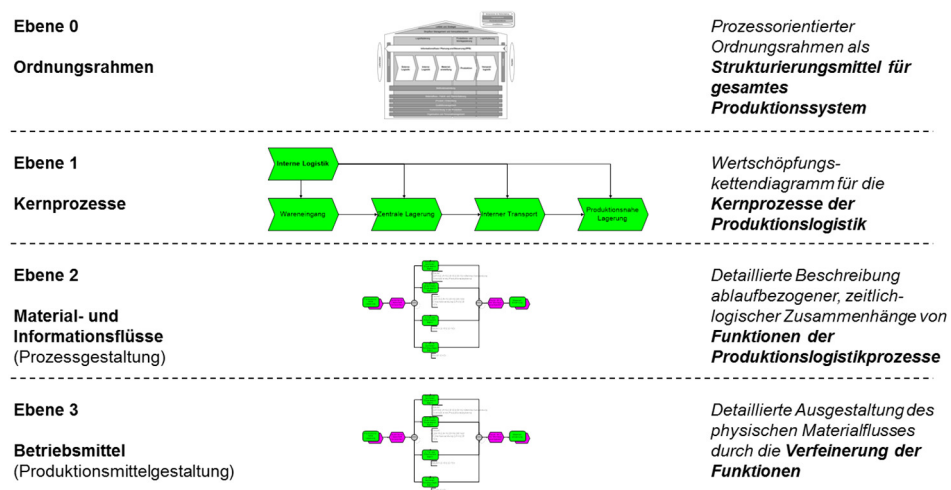
Bei der Anwendergruppe des Referenzmodells handelt es sich vor allem um Fachplaner der Produktionslogistik in KMUs. Somit ist die beschriebene Referenzmodellkonstruktion im Sinne von ARIS auf der Beschreibungsebene des Fachkonzeptes einzuordnen (Scheer 2002). Dementsprechend wird vorrangig die Ereignisgesteuerte Prozesskette (EPK) als semi-formale Modellierungssprache verwendet.

Die Entwicklung des Referenzmodelles basiert einerseits auf der deduktiven Ableitung idealtypischer Prozesse aus den theoretischen Ausarbeitungen zu Lean Production und ganzheitlichen Produktionssystemen. Andererseits erfolgt die individuelle Identifikation von Prozesslösungen und Adaptionparametern durch die praktische Produktionslogistikgestaltung bei einem mittelständischen Unternehmen (Induktion). Somit ergibt sich eine praxisgestützte Konstruktion, welche auf einer theoretischen Fundierung basiert und die unmittelbaren praktischen Aufgabenstellungen berücksichtigt.

### 5.1 Ordnungsrahmen und Ebenen des Referenzmodells

Aufgrund der hohen Komplexität im Rahmen der Prozessplanung in Produktionssystemen ist zu Beginn ein geeigneter Ordnungsrahmen als Systembegrenzung und Strukturierungsmittel zu entwickeln (Meise 2001; Schubel et al. 2015a). Zusätzlich besteht das Referenzmodell aus mehreren Detaillierungsebenen, welche die wesentlichen Kernprozesse und Funktionen des Materialflusses beinhalten und diese bis hin zur Betriebsmittelauswahl verfeinern (Bild 2).





**Bild 2: Ebenen des Referenzmodells für die Produktionslogistikgestaltung**

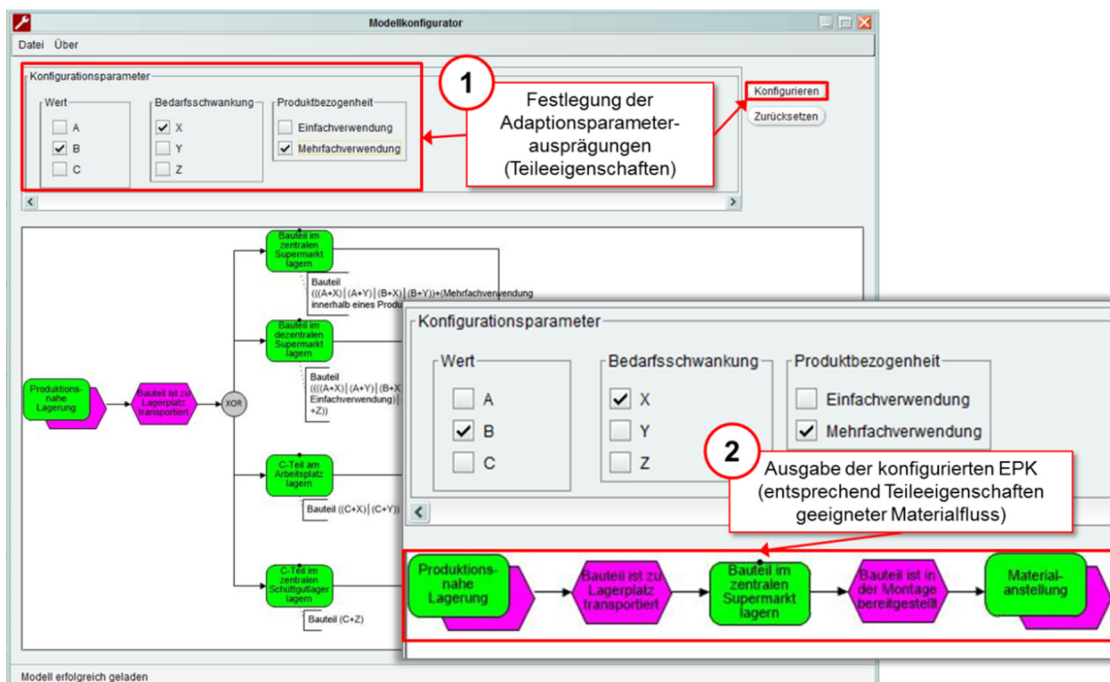
## 5.2 Gestaltung des physischen Materialflusses anhand konfigurierbarer EPKs

Der Ordnungsrahmen (Ebene 0) und die Kernprozesse (Ebene 1) des Referenzmodells ermöglichen einen gezielten Einstieg in die konkrete Gestaltung des Materialflusses auf Ebene 2. Anhand konfigurierbarer EPKs erfolgt dort die situative Gestaltung des Materialflusses durch die Auswahl der benötigten Funktionen. Der dafür notwendige Adaptionsmechanismus nach (Delfmann 2006) basiert auf Konfigurationstermen und Adaptionsparametern. Die Adaptionsparameter umfassen für die Gestaltung des Materialflusses der Produktionslogistik vor allem Teileigenschaften (z.B. Wert, Bedarfsschwankung und Produktbezogenheit des Bauteils). Die Verfeinerung der benötigten Materialflussfunktionen durch die Produktionsmittelgestaltung ist auf Ebene 3 vorgesehen.

Zur Veranschaulichung der Gestaltung des physischen Materialflusses anhand des adaptiven Referenzmodelles soll ein konkretes Beispiel dienen: Es gilt den physischen Materialfluss für ein Bauteil im Rahmen der *internen Logistik* (Ebene 0) in der *produktionsnahen Lagerung* (Ebene 1) zu gestalten. Für das Bauteil wird, entsprechend dessen Adaptionsparameterausprägungen, die Funktion eines *Dezentralen Supermarktes* benötigt. Die Funktion des *Dezentralen Supermarktes* ist auf Ebene 3 durch Auswahl der genauen Lagerform und -hilfsmittel zu verfeinern. Das Ergebnis ist der konkrete Materialfluss eines spezifischen Bauteils in der produktionsnahen Lagerung. Abschließend gilt es für den Logistikplaner die Prozessplanungen durch das Einbeziehen der Rahmenbedingungen, welche in den Adaptionsparametern nicht berücksichtigten sind, zu finalisieren.

## 5.3 Umsetzung in einem Softwarewerkzeug

Die Abbildung des Referenzmodells erfolgt softwaregestützt. Der im Zuge der Forschungsarbeiten entwickelte Software-Prototyp ermöglicht dem Logistikplaner die Auswahl von Adaptionsparameterausprägungen entsprechend des vorliegenden Planungsfalls. Daraufhin wird dem Logistikplaner der situativ angepasste Prozess ausgegeben, welcher als Basis für die abschließenden Feinplanungen bei der Materialflussgestaltung dient (Bild 3).



**Bild 3: Abbildung des Referenzmodelles durch einen Software-Prototyp**

## 6 Fazit und Ausblick

Die ersten praktischen Erfahrungen zeigen, dass das *adaptive Referenzmodell für die Materialflussgestaltung in der Produktionslogistik* eine effiziente und effektive Prozessplanung ermöglicht. Vor allem die Abstimmungsaufwände im Rahmen der Materialflussplanung werden reduziert, da als Standard festgelegte Materialflüsse und Adaptionparameter kürzere und weniger Planungsrunden im Zuge der teilebezogenen Materialflussgestaltung erfordern. Als Planungskonzept verstanden, kann das adaptive Referenzmodell zudem zur konkreten Umsetzung einer flexiblen Prozessstandardisierung in der Produktionslogistik genutzt werden.

In den zukünftigen Forschungsarbeiten gilt es die adaptiven Referenzprozesse für die Materialbereitstellung in der Produktionslogistik zu vervollständigen und auf Ebene 3 zu verfeinern. Im Zuge dessen sind weiterhin die geeigneten Adaptionparameter für die Gestaltung der Produktionslogistik in KMUs zu identifizieren. Abschließend wird das Konzept anhand praktischer Anforderungen und Planungsfälle evaluiert.

## 7 Literatur

- Bauernhansl T, Hompel M ten, Vogel-Heuser B (2014) Industrie 4.0 in Produktion, Automatisierung und Logistik. Springer Fachmedien Wiesbaden, Wiesbaden
- Becker J, Schütte R (2004) Handelsinformationssysteme: Domänenorientierte Einführung in die Wirtschaftsinformatik, 1. Auflage. mi-Wirtschaftsbuch, München
- Binner HF (2010) Handbuch der prozessorientierten Arbeitsorganisation: Methoden und Werkzeuge zur Umsetzung, 4. Auflage. Hanser, München
- Delfmann P (2006) Adaptive Referenzmodellierung: Methodische Konzepte zur Konstruktion und Anwendung wiederverwendungsorientierter Informationsmodelle. Logos-Verl, Berlin

- Fettke P, Loos P (2004) Referenzmodellierungsforschung: Langfassung eines Aufsatzes. Working Papers of the Research Group ISYM(16)
- Fettke P, Loos P (2005) Der Beitrag der Referenzmodellierung zum Business Engineering. HMD - Praxis der Wirtschaftsinformatik 42(241):18–26
- Günthner WA, Boppert J (2013) Lean Logistics: Methodisches Vorgehen und praktische Anwendung in der Automobilindustrie. Springer, Berlin, Heidelberg
- Hevner A, Chatterjee S (2010) Design Research in Information Systems, vol 22. Springer US, Boston, MA
- Krebs M, Goßmann D, Erohin O, Bertsch S, Deuse J, Nyhuis P (2011) Standardisierung im wandlungsfähigen Produktionssystem: Einfluss der Prozess- und Ressourcenstandardisierung auf die Wandlungsfähigkeit. ZWF 106(12):912–917
- Meise V (2001) Ordnungsrahmen zur prozessorientierten Organisationsgestaltung: Modelle für das Management komplexer Reorganisationsprojekte. Schriftenreihe Studien zur Wirtschaftsinformatik, vol 10. Kovač, Hamburg
- Pawellek G (2007) Produktionslogistik: Planung - Steuerung - Controlling : mit 42 Übungsfragen. Hanser, München
- Scheer AW (2002) ARIS: vom Geschäftsprozess zum Anwendungssystem, 4. Aufl.. Springer, Berlin
- Schmelzer HJ, Sesselmann W (2013) Geschäftsprozessmanagement in der Praxis: Kunden zufriedenstellen, Produktivität steigern, Wert erhöhen; [das Standardwerk], 8., überarb. und erw. Aufl. Hanser, München
- Schubel A, Seel C, Schneider M (2015a) Ein Ordnungsrahmen für die Produktions- und Logistikprozessplanung in kleinen und mittelständischen Unternehmen mit diskreter Produktion. In: Barton T, Erdlenbruch B, Herrmann F, Müller C, Marfurt K, Seel C (eds) Angewandte Forschung in der Wirtschaftsinformatik. mana-Buch, Heide, pp 48–57
- Schubel A, Seel C, Schneider M (2015b) Informationsmodelle für die Produktions- und Logistikplanung: Eine Literaturanalyse des aktuellen Referenzmodellbestands. In: Thomas O, Teuteberg F (eds) 12. Internationale Tagung Wirtschaftsinformatik, Osnabrück, pp 527–541
- Schubel A, Seel C, Schneider M (2015c) Informationsmodellierung in der Produktions- und Logistikplanung: Ein Abgleich des aktuellen Referenzmodellebestandes mit den praktischen Anforderungen in der Produktionslogistik von kleinen und mittelständischen Unternehmen. Landshuter Arbeitsberichte zur Wirtschaftsinformatik, Landshut
- Staiger M (2008) Wissensmanagement in kleinen und mittelständischen Unternehmen: Systematische Gestaltung einer wissensorientierten Organisationsstruktur und -kultur, 1. Aufl. Weiterbildung, Bd. 6. Rainer Hampp Verlag, Mering
- Thomas O, Scheer A (2006) Business Engineering mit Referenzmodellen: Konzeption und informationstechnische Umsetzung. Information Management & Consulting 21(1):65–71
- Westkämper E (2013) Struktureller Wandel durch Megatrends. In: Westkämper E, Spath D, Constantinescu C, Lentjes J (eds) Digitale Produktion. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 7–9



# Didaktisches Planspiel zum Customizing-Prozess

Jens Siewert<sup>1</sup> und Stefan Wunderlich<sup>1</sup>

<sup>1</sup> Carl von Ossietzky Universität Oldenburg, Abteilung Wirtschaftsinformatik,  
{jens.siewert | stefan.wunderlich}@uni-oldenburg.de

## Abstract

Die Herausforderungen und Abläufe bei der Einführung von Standardsoftware in Unternehmen sollen Studierenden in einer Übung begleitend zur Vorlesung „Customizing“ praxisorientiert vermittelt werden. Hierfür wurde ein Planspiel entworfen, das ein möglichst realistisches Customizing-Szenario abbildet. Dabei wird der didaktische Fokus nicht nur auf die Nutzung und Anwendung von Standardsoftware gelegt, sondern darüber hinaus die Organisations- und Prozesssicht einbezogen. Die Übung wird am SAP-ERP-System als Gruppenarbeit durchgeführt und beinhaltet sowohl die Analyse von Unternehmensstrukturen und -abläufen, als auch das Customizing der Standardsoftware. Im Rahmen des Planspiels werden sämtliche Lernstufen nach Bloom adressiert. Die Studierenden sollen Inhalte nicht nur nachvollziehen, sondern darüber hinaus verstehen und selbstständig anwenden. Durch die Teilnahme am Planspiel vertiefen die Studierenden unabhängig von ihren Vorerfahrungen, neben den Abläufen im Customizing-Prozess, auch den eigenständigen Umgang mit der Standardsoftware. Eine ganzheitliche didaktische Betrachtung des Customizing-Prozesses wird auf diese Weise realisiert.

## 1 Einleitung

Customizing wird als Anpassen von Standardsoftware an die Anforderungen eines Kunden verstanden (Mertens 1997). Dabei wird zwischen Customizing im engeren Sinne und Customizing im weiteren Sinne unterschieden. Customizing im engeren Sinne wird dabei als konkrete Anpassung der Software an den Anwendungsfall verstanden, z. B. Aufgaben der Modularisierung bzw. Parametrisierung von Standardsoftware (Stahlknecht und Hasenkamp 2002). Des Weiteren erlaubt Customizing die Erweiterung der Software, beispielsweise vom Softwarehersteller vorgesehene Erweiterungen, Modifikationen am Standard-Repository oder komplette Eigenentwicklungen (Frick et al. 2008).

Im Rahmen der Übung zur Begleitung der Vorlesung „Customizing“ des Studienganges M. Sc. „Wirtschaftsinformatik“ werden den Studierenden sowohl die Benutzung der Standardsoftware, als auch die Aufgaben und Probleme des Customizing vermittelt. In diesem Paper werden die Aufgaben und Probleme des Customizing beschrieben. Anschließend wird die Ausgangssituation dargelegt, die der Grund für die Entwicklung der neuen Übung gewesen ist. Im Anschluss wird der Aufbau des Planspielkonzepts (Meyer 1998) erläutert, das aktuell in der Übung umgesetzt wird.

## 2 Herausforderungen des Customizing

Neben dem Customizing im engeren Sinne, das hauptsächlich Kernaufgaben der Wirtschaftsinformatik umfasst, ergeben sich gerade im Customizing im weiteren Sinne ungewohnte Herausforderungen sowohl für das Unternehmen, als auch für die handelnden Personen.

Es muss eine finanzielle Belastung bewältigt werden (Gadatsch 2010), die Prozesse des Unternehmens müssen auf die Software übertragen werden und gleichzeitig muss die zusätzliche Arbeitslast abgefangen werden (Hansen et al. 2015). Auch müssen die Mitarbeiter des Unternehmens von den bevorstehenden Veränderungen überzeugt werden (Bingi et al. 1999). Deshalb sollte ein IT-Berater, der ein Customizing-Projekt begleitet, je nach Aufgabenstellung Expertise in folgenden Themengebieten besitzen (Berufenet 2015):

- betriebliche IT-Prozesse,
- profilspezifische IT-Fachaufgaben und
- Mitarbeiterführung und Personalmanagement.

## 3 Ausgangssituation

Im Rahmen der M. Sc. „Wirtschaftsinformatik“ Lehrveranstaltung „Customizing“ (6 ECTS, 2+2 SWS) sollte eine passende Übung kreiert werden, die den Studierenden die theoretisch-deskriptiven Vorlesungsinhalte praxisorientiert vermittelt. Diese ergeben sich aus den oben beschriebenen Aufgaben für IT-Berater und sind unter anderem:

- betriebliche Organisation,
- Prozessanalyse und -modellierung,
- Einführungsmethoden und
- Change Management.

Die zuvor verwendeten Übungsmaterialien bestanden lediglich aus Fallstudien (Weidner 2008) zu verschiedenen betrieblichen Anwendungssystemen, sowie einzelnen Übungen zum Thema Geschäftsprozessmodellierung bzw. Workflowmanagement ohne ganzheitlichen Charakter. Durch den Aufbau der Fallstudien wurden die Studierenden lediglich zum Verstehen von Wissen angeregt. Die weiteren Lernstufen nach Bloom (Bloom et al. 1956), also das Durchdenken eines Problems bzw. das selbstständige Anwenden des Wissens, wurden nur selten erreicht (Peters et al. 2012).

Dies begründet sich insbesondere durch den speziellen Hintergrund der Übungsteilnehmer. Der Kurs wird gezielt für Studierende des Studienganges M. Sc. „Wirtschaftsinformatik“ angeboten. Man kann den Teilnehmern damit nicht nur eine grundlegende IT-Affinität unterstellen, sondern die meisten bringen bereits erste Erfahrung im Umgang mit SAP-Produkten aus dem Bachelorstudium oder der Praxis mit (BC 2015).

Da zudem die Customizing-Fallstudie des SAP University Competence Centers (UCC) (Ebner et al. 2009) in Magdeburg am aktuell im Lehrstuhl eingesetzten System (SAP ERP ECC 6.04 IDES, SAP ERP ECC 5.0 benötigt) nicht mehr zur Verfügung steht bzw. nur noch eingeschränkt auf diesem System nutzbar ist, war es erforderlich eine neue Übung zu konstruieren. Die einzelnen Übungsaufgaben wurden in ein ganzheitlich strukturiertes Planspiel, das die Herausforderungen

eines Customizing-Szenarios widerspiegelt und durch eine möglichst freie Gestaltung den Lernerfolg der Studierenden maximiert, übertragen.

Im Studiengang M. Sc. „Wirtschaftsinformatik“ an der Universität Oldenburg wird den Studierenden zur Motivation die Möglichkeit gewährt, sich durch die Übung einen Bonus für die Klausur zu erarbeiten. Sie erhalten für bestandene Zwischenziele je nach Schwierigkeitsgrad und Umfang einen oder mehrere Punkte als Bonus. Insgesamt können Sie bei bestandener Klausur bis zu zehn zusätzliche Prozentpunkte erhalten.

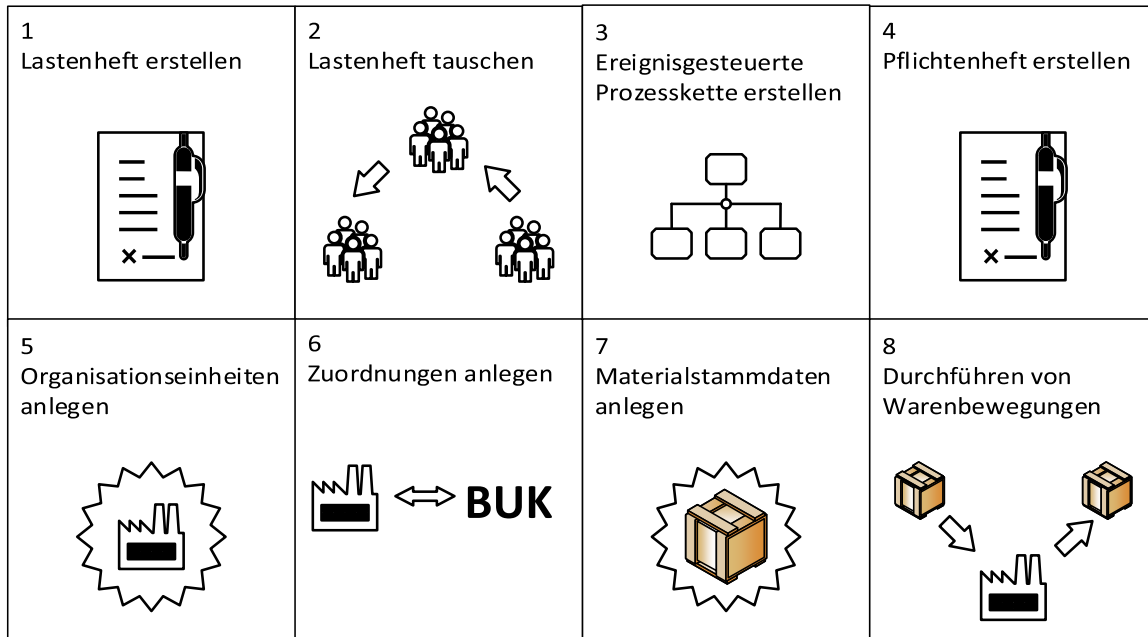
## 4 Aufbau des Planspiels

Ziel des Planspiels ist die Simulation eines möglichst vollständigen und realistischen Customizing-Szenarios. Dabei arbeiten sich die Studierenden tiefer in die Themen Geschäftsprozessmodellierung, Lasten- und Pflichtenheft sowie Customizing im engeren Sinne ein. Als Ergebnis der Übung sollen Buchungen für ein selbst konzipiert und eingepflegtes Unternehmen am SAP ERP ECC 6.04 IDES System durchführbar sein. Dabei haben die Studierenden parallel zu den nicht mit der Standardsoftware durchzuführenden Komponenten des Planspiels die Möglichkeit sich mit dem SAP-ERP-System vertraut zu machen. Dafür stehen die folgenden Fallstudien des SAP UCC (Weidner 2008) zur Verfügung:

- **Navigationskurs SAP ERP 6.0 UCC:** Die Navigation des SAP-Systems wird vermittelt. Vorkenntnisse sind nicht erforderlich.
- **Customizing-Fallstudie SAP ERP ECC 5.0:** Ein fiktives Unternehmen wird vorgestellt und dessen Aufbauorganisation rudimentär im SAP-ERP-System abgebildet.
- **Integrations-Fallstudie PP SAP UCC:** Materialstammdaten werden angelegt und im Anschluss in einem einfachen Produktionsprozess verwendet.
- **Integrations-Fallstudie Logistik SAP UCC:** Es wird die Integration von drei Modulen aus der Logistik aufgezeigt. Im Anschluss wird der Zusammenhang zum Modul Finanzwesen hergestellt. Dabei werden weitere Stammdaten eingeführt.
- **Integrations-Fallstudie CO SAP UCC:** Die Kostenstellenrechnung mit ihren wesentlichen Stammdaten und Planungsmethoden wird bearbeitet. Abschließend erfolgt eine Integration der Kostenstellenrechnung in die Produktionsplanung.

Durch die Bearbeitung dieser Fallstudien erhalten die Studierenden bereits einen Einblick in die erforderlichen Transaktionen und die Umgebung im SAP-ERP-System.

Das Planspiel selbst gliedert sich in acht Zwischenziele. Durch kurze, klare Fristen für jedes Zwischenziel wird sichergestellt, dass die Studierenden sich regelmäßig mit der Übung beschäftigen. Die Zwischenziele sind in Bild 1 dargestellt und werden im Folgenden ausführlich erläutert.



**Bild 1: Zwischenziele des Planspiels, Quelle: eigene Darstellung**

Zu Beginn des Planspiels bilden die Kursteilnehmer Übungsgruppen zu je drei Personen. Jede Dreiergruppe erarbeitet selbstständig ein virtuelles Unternehmen (VU). Die Informationen, die die Gruppen zu den VU zusammenstellen, werden in einem Lastenheft gebündelt. Für das Lastenheft werden diese in beschreibender Form (Prosa) dokumentiert (Zwischenziel 1). Bedingt durch den begrenzten Zeitrahmen der Übung von 2 SWS ist es wichtig, dass die VU weder zu komplex noch zu kompliziert werden. Dennoch ist es essenziell, gezielte Informationen zu den jeweiligen VU einzufordern. Zu diesem Zweck werden von der Kursleitung konkrete Anforderungen (u. a. Branche, Organisationsstruktur, Kerngeschäftsprozesse, Material- und Mitarbeiterstammdaten) an die VU vorgegeben.

Die Lastenhefte werden nach Fertigstellung nach Rotationsprinzip zwischen den Gruppen weitergegeben (Zwischenziel 2). Diese bilden in der Übung jeweils die Grundlage für eine andere Übungsgruppe. Ziel dieser Maßnahme ist es, eine möglichst realistische Customizing-Situation nachzustellen. Customizing-Projekte mit SAP-Produkten werden oftmals von externen Dienstleistern durchgeführt (Gadatsch 2010). Obwohl der externe Dienstleister das Unternehmen zu Beginn des Customizing-Projektes noch nicht kennt, soll er dessen Geschäftsprozesse und Organisationsstrukturen in möglichst kurzer Zeit innerhalb der Standardsoftware abbilden. Der Tausch der Lastenhefte versetzt die Studierenden in die Situation eines externen Dienstleisters. Sie kennen das Unternehmen, das sie innerhalb der Standardsoftware abbilden sollen zu Beginn nicht, sollen dieses jedoch sehr zeitnah umsetzen. Die dafür benötigten Informationen erhalten die Studierenden aus dem an sie übergebenen Lastenheft, sowie durch Kommunikation mit der Gruppe, die ihr das Lastenheft übergeben hat. Dadurch steht jede Gruppe mit zwei weiteren Gruppen in Kontakt. Die Gruppen müssen sich auf gegenseitige Anfrage mit weiteren Informationen versorgen. Durch dieses Vorgehen werden zusätzlich die Kommunikationsprozesse (Nah et al. 2001) zwischen Auftraggeber und Auftragnehmer innerhalb eines Customizing-Projekts abgebildet.



Zwischenziel 3 ist für die Gruppen die Erstellung einer Modellierung zweier Kerngeschäftsprozesse des VUs. Die Modellierung erfolgt dabei als Ereignisgesteuerte Prozesskette (EPK). Die Vorgabe ergibt sich aus der Notwendigkeit der Vergleichbarkeit sowie aus den Vorlesungsinhalten, deren Bestandteil die Prozessmodellierung mit EPK ist. Als Kernprozesse wurden der Einkaufsprozess sowie der Verkaufsprozess festgelegt. Anschließend erstellt jede Gruppe ein Pflichtenheft für das Customizing-Vorhaben (Zwischenziel 4). Dabei soll das Pflichtenheft die folgenden wesentlichen Informationen enthalten.

Einen *Meilensteinplan* mit Informationen bis wann die Aufbauorganisation im System hinterlegt sein soll und zu welchem Termin die Ablauforganisation im System hinterlegt ist. Weiter soll dargestellt werden, aus welchen Organisationseinheiten die *Aufbauorganisation* besteht und wie diese in das System übernommen werden. Durch die vorher erstellten EPK wird die *Ablauforganisation* in das Pflichtenheft aufgenommen. Zusätzlich wird benannt, welche Organisationseinheiten bei den Kerngeschäftsprozessen beteiligt sind.

Die Modellierung der Prozesse ist Teil des Pflichtenheftes, da dies der Struktur eines Pflichtenheftes nach Gronau (Gronau 2001) entspricht. Durch das Anfertigen des Pflichtenheftes durchdringen die die Struktur der jeweiligen VU, sodass sie im Anschluss die Kompetenz besitzen, das VU mit seinen Prozessen im SAP-System einzupflegen.

Geschult durch das Bearbeiten der zur Verfügung gestellten Fallstudien übertragen die Studierenden im Hauptteil des Planspiels das VU, dessen Beschreibung sie im Pflichtenheft erstellt haben, in das SAP-ERP-System. Dies bedeutet, dass sie in der Standardsoftware die folgenden Arbeitsschritte durchlaufen:

Zu Beginn legen die Studierenden analog zur Customizing-Fallstudie die Organisationseinheiten der VU an (Zwischenziel 5). Dafür begeben sie sich in den SAP Reference Implementation Guide (IMG) (Gronau 2001). Danach werden ebenfalls im SAP Reference IMG die Zuordnungen zwischen den Organisationseinheiten angelegt (Zwischenziel 6). Da die von den Studierenden erstellten VU nicht dem Beispielunternehmen in der Customizing Fallstudie entsprechen, ist eine bedeutende Transferleistung der Studierenden erforderlich.

Nachdem die Studierenden die Organisationseinheiten im SAP Reference IMG angelegt haben, können Sie im nächsten Schritt die Materialstammdaten des VUs anlegen (Zwischenziel 7). Dies wird ebenfalls beispielhaft in der Integrationsfallstudie ausgeführt, muss von den Studierenden aber auf die spezifische Situation übertragen werden.

Nach dem Anlegen der Materialstammdaten folgt das abschließende Zwischenziel 8 (die Durchführung von Warenbewegungen). Entsprechend der Vorgaben im Lasten- und Pflichtenheft sollen die Studierenden sowohl einen Wareneingang, als auch einen Warenausgang virtuell realisieren. Hierbei können sich die Studierenden erneut an den Vorgaben der Fallstudien orientieren. Es ist allerdings zu beachten, dass diese Aufgabe deutlich über die Anweisungen aus den Fallstudien hinausgeht. Um das Planspiel erfolgreich zu beenden, müssen die Studierenden anhand der Fehlermeldungen des Systems eigenständig auf bestehende Abweichungen zwischen Fallstudien und Planspiel reagieren und selbstständig Lösungen erarbeiten. Dazu können sie auf das Internet und insbesondere auf das SAP Help Portal (Frick et al. 2008) zurückgreifen. Dies erfordert ein intensives Auseinandersetzen mit dem System und den VU.

## 5 Evolution des Planspiels

Das Planspiel wurde bereits zum zweiten Mal im Rahmen der Vorlesung „Customizing“ durchgeführt. Erstmals wurde das Planspiel im Sommersemester 2014 eingesetzt. Im Rahmen dieser ersten praktischen Umsetzung des Planspiels wurden lediglich zwei Gruppen gebildet. Dementsprechend wurden von den Studierenden parallel zwei VU in das SAP-ERP-System eingepflegt. Die höhere Gruppengröße von 15 Personen sollte dazu führen, dass jeder Studierende jeweils eine spezielle Rolle im VU übernimmt. Jeder Studierende sollte einen Experten im Unternehmen verkörpern, der für einen spezifischen Bereich verantwortlich ist. Gleichzeitig sollte jeder dieser Experten der exklusive Ansprechpartner für seinen Bereich sein. Da für viele Studierende trotz Vorerfahrung im Umgang mit dem SAP-ERP-System das Customizing des Systems eine neue Herausforderung darstellte, traten die exakten Anpassungen der Software an die Prozesse des VUs deutlich in den Hintergrund. Da auf diese Weise das exklusive Expertenwissen nicht erforderlich war, konnten hierdurch auch die Expertenrollen nicht ausgefüllt werden.

Da die Größe der Gruppen dafür gesorgt hat, dass einzelne Studierende den Hauptteil des Customizing im engeren Sinne übernommen haben und ein Großteil der Gruppe dabei lediglich unterstützend fungierte, fiel für das kommende Jahr die Entscheidung, die Gruppengröße zu reduzieren. Die Gruppengröße von aktuell drei Studierenden gibt den Studierenden weiterhin die Möglichkeit sich auszutauschen und kreativ tätig zu werden.

Im Rahmen der zweiten Durchführung des Planspiels haben sich die Veränderungsmaßnahmen als Erfolg erwiesen (Siewert 2015). Lediglich der Aspekt des Change Managements ist, durch die Reduzierung der Gruppengröße in den Hintergrund getreten. Weiterhin konnte festgestellt werden, dass bei vielen Studierenden während des Customizings das Pflichtenheft und insbesondere die Modellierung aus dem Fokus gerieten. Dies ist darauf zurückzuführen, dass die Studierenden während der Übung häufig erstmalig mit dem Customizing einer Standardsoftware konfrontiert werden. Auch wenn diese Tatsache nicht den Lernerfolg gefährdet, werden die Studierenden, die an der nächsten Durchführung des Planspiels beteiligt sind, zum Ende aufgefordert werden, einen Soll-Ist Abgleich zwischen ihren eigenen Vorgaben und dem finalen Ergebnis zu erstellen. Auf diese Weise soll eine Reflektion des eigenen Vorgehens erzielt werden.

## 6 Fazit und Ausblick

In diesem Paper wurden die Möglichkeiten des Einsatzes eines Planspiels in der Lehre aufgezeigt. Für angehende Wirtschaftsinformatiker, die sich mit Customizing in der Arbeitswelt auseinandersetzen werden, wird hier praxisnah Wissen zu den Themen (I) Instrumente des Customizing, wie z. B. der Einsatz von Lasten- und Pflichtenheft, (II) Herausforderungen des Customizing-Prozesses, (III) Aufbau und Handhabung von Standardsoftware am Beispiel von ERP-Systemen, (IV) Sozialkompetenz, z. B. durch Kommunikation in und zwischen Gruppen und (V) Praxiserfahrung am SAP-ERP-System vermittelt. Durch die Konzeptionierung als Gruppenübung mit freier Gestaltung wird ein hoher Lernerfolg erzielt (BC 2015). Bemerkenswert ist, dass sich der Lernerfolg einstellt, unabhängig davon, ob Vorerfahrungen mit Standardsoftware bestanden. Dies ist darauf zurück zu führen, dass durch das Planspiel alle Lernstufen nach Bloom (Bloom et al. 1956) adressiert werden. Einige Aspekte des Customizing-Prozesses werden allerdings noch nicht berücksichtigt und sollen zukünftig schrittweise in das Planspiel eingebunden werden. Dazu zählen insbesondere die Risiken eines Customizing-Prozesses. Die Übertragbarkeit des Planspiels auf andere Standardsoftwaresysteme ist dagegen grundsätzlich möglich, dafür ist ein Ersatz der durch

die begleitend zur Übung bearbeiteten Fallstudien erforderlich. Durch den großen didaktischen Erfolg des Planspiels wurde die Entscheidung getroffen, diese Übung auch in den kommenden Lehrveranstaltungen „Customizing“ einzusetzen.

## 7 Literatur

- BERUFENET (2015) Steckbrief IT-Berater/in. <http://berufenet.arbeitsagentur.de/berufe/docroot/r2/blobs/pdf/bkb/14040.pdf>. Abgerufen am 30.07.2015
- Bloom BS, Engelhart MD, Furst EJ, Hill WH, Krathwohl DR (1956) Taxonomy of educational objectives – The Classification of Educational Goals Handbook I: Cognitive Domain. David McKay Company, Inc., New York
- Bingi P, Sharma MK, Godla JK (1999) Critical Issues Affecting an ERP Implementation. *Information Systems Management*. 16(3):7-8
- Ebner W, Leimeister JM, Krcmar H (2009) Community engineering for innovations: the ideas competition as a method to nurture a virtual community for innovations. *R&D Management*. 39(4):342–356
- Frick D, Gadatsch A, Schäffer-Kütz UG (2008) Grundkurs SAP ERP Geschäftsprozessorientierte Einführung mit durchgehendem Fallbeispiel. Friedr. Vieweg & Sohn, Wiesbaden
- Gadatsch A (2010) Grundkurs Geschäftsprozess- Management Methoden und Werkzeuge für die IT-Praxis: Eine Einführung für Studenten und Praktiker. 6. Auflage. Vieweg + Teubner, Wiesbaden
- Gronau N (2001) Industrielle Standardsoftware Auswahl und Einführung. Oldenbourg, München
- Hansen HR, Mendling J, Neumann G (2015) Wirtschaftsinformatik. 11. Auflage. De Gruyter, Berlin
- Mertens P (1997) Lexikon der Wirtschaftsinformatik. 3. Auflage. Springer, Berlin
- Meyer H (1989) Unterrichtsmethoden 2. Praxisband. 2. Auflage. Cornelsen, Berlin
- Nah FF, Lau JL, Kuang J (2001) Critical factors for successful implementation of enterprise systems. *Business Process Management Journal*. 7(3):285 – 296
- Peters D, Haak L, Marx Gómez J, (2012) Adaptive Learning Cycle to Improve the Competence-Building for Enterprise Systems in Higher Education. In: Varajão J Cruz-Cunha M M Trigo A (Hrsg.) *Organizational Integration of Enterprise Systems and Resources: Advancements and Applications*. IGI Global, Hershey, 76-99
- Siewert J (2015) Evaluation eines Customizing Planspiels in der Veranstaltung “Customizing” im Studiengang M. Sc. Wirtschaftsinformatik. [https://www.uni-oldenburg.de/fileadmin/user\\_upload/informatik/ag/vlba/download/Customizing\\_Evaluation\\_Uebung\\_SoSe15.pdf](https://www.uni-oldenburg.de/fileadmin/user_upload/informatik/ag/vlba/download/Customizing_Evaluation_Uebung_SoSe15.pdf). Abgerufen am 30.07.2015
- Stahlknecht P, Hasenkamp U (2002) Einführung in die Wirtschaftsinformatik. 10. Auflage. Springer, Berlin
- Weidner S (2008) Fallstudien SAP ERP. <https://portal.ucc.uni-magdeburg.de/irj/portal/anonymous/login?NavigationTarget=navurl://37ae857d98e1e89226acf26002262ad2&LightDTNKnobID=-1077532283&windowId=WID1438268168498>. Abgerufen am 30.07.2015



# Vergleichende Betrachtung von Software-Werkzeugen zur Geschäftsmodellentwicklung

Sebastian Strotmeyer<sup>1</sup>, Thomas John<sup>2</sup> und Dennis Kundisch<sup>3</sup>

<sup>1</sup> Universität Paderborn, stroti@mail.uni-paderborn.de

<sup>2</sup> Universität Paderborn, Lehrstuhl für Wirtschaftsinformatik, Information Management & E-Finance, C-LAB, thomas.john@uni-paderborn.de

<sup>3</sup> Universität Paderborn, Lehrstuhl für Wirtschaftsinformatik, Information Management & E-Finance, Dennis.Kundisch@wiwi.uni-paderborn.de

## Abstract

Die Relevanz von Geschäftsmodellen hat enorm zugenommen und als Basis für die Geschäftsmodellentwicklung hat sich insbesondere die Business Model Canvas etabliert. Zur Unterstützung der Canvas-basierten Geschäftsmodellentwicklung wurden auch bereits zahlreiche Software-Werkzeuge vorgeschlagen. Derzeit ist jedoch unklar, welche Funktionen diese Klasse von Werkzeugen grundsätzlich besitzt und anhand welcher Eigenschaften potenzielle Nutzer ein Werkzeug auswählen sollten. Daher wird in dieser Arbeit ein umfassender Überblick über aktuell verfügbare Software-Werkzeuge zur Canvas-basierten Geschäftsmodellentwicklung gegeben und basierend auf 22 identifizierten Werkzeugen eine Taxonomie typischer Eigenschaften entwickelt. Diese Taxonomie erleichtert Praktikern die Auswahl eines für sie geeigneten Werkzeugs und kann für Forscher eine Basis für kumulative Forschung zur Softwareunterstützung von Geschäftsmodellentwicklungsprozessen darstellen.

## 1 Einleitung

Ein Geschäftsmodell beschreibt die grundlegenden Mechanismen, mit denen ein Unternehmen Werte für Kunden erzeugt und damit Gewinn erzielt (Teece 2010). Das Interesse an Geschäftsmodellen hat in den letzten Jahren in Forschung und Praxis rasant zugenommen. Beispielsweise zeigt eine aktuelle Studie in insgesamt 26 Ländern, dass 60 % der 3000 befragten Manager es im Rahmen ihrer Innovationsaktivitäten für eine große Herausforderung halten, effektive Geschäftsmodelle zu entwickeln (GE 2014). Laut einer IBM-Studie unter mehr als 1000 CEOs halten praktisch alle CEOs Geschäftsmodellinnovationen für grundsätzlich wünschenswert – und mehr als zwei Drittel streben sogar weitreichende (“extensive“) Geschäftsmodellinnovationen an (IBM 2008, S. 48).

Während in der Forschung noch umstritten ist, welches Geschäftsmodellverständnis bzw. welche Definition eines Geschäftsmodells sinnvollerweise zu wählen ist (Zott et al. 2011), hat sich in der unternehmerischen und universitären Praxis mit dem *Business Model Canvas* (Osterwalder und

Pigneur 2010) bereits ein Standard etabliert. Das zugehörige Buch wurde bereits rund eine Million Mal verkauft (Strategyzer 2015) und es gibt zahlreiche Multiplikatoren, die zur weiteren Verbreitung des Canvas in der Praxis beitragen. So liefert beispielsweise eine Google-Suche nach *workshop AND "business model canvas"* mehr als 100.000 Treffer<sup>1</sup>.

Während die Canvas als konzeptioneller Rahmen für die Geschäftsmodellentwicklung etabliert ist, ist noch weitgehend unklar, durch welche zusätzlichen Hilfestellungen insbesondere die kreative erste Phase der Geschäftsmodell-Entwicklung mit dem Canvas unterstützt werden kann. Software-basierten Werkzeugen wird ein großes Potenzial zugeschrieben, um kreative Prozesse im Allgemeinen (Shneiderman 2007) und Geschäftsmodellentwicklungsprozesse im Speziellen (Osterwalder und Pigneur 2013) zu unterstützen. Es wurden auch bereits einige solche Werkzeuge in der wissenschaftlichen Literatur (z. B. Business Model Developer) und in der Praxis (z. B. Canvanizer, Business Model Toolbox) vorgestellt. Unklar ist allerdings, welche Funktionen und weiteren Eigenschaften für diese neuartige Klasse von Software-Werkzeuge charakteristisch sind und welche Kriterien potenziellen Nutzern dabei helfen können, das für sie geeignete Werkzeug zur Geschäftsmodellentwicklung zu identifizieren. Dadurch wird die Verbreitung dieser Werkzeuge in der Praxis gehemmt und der wissenschaftliche Fortschritt erschwert, da eine geeignete Basis für kumulative Forschung fehlt. Daher ist es das Ziel dieser Arbeit, einen umfassenden Überblick über aktuell verfügbare Software-Werkzeuge zur Canvas-basierten Geschäftsmodellentwicklung zu erarbeiten. Dafür werden durch eine umfassende Recherche entsprechend der Methode von Wolfswinkel et al. (2013) die aktuell im wissenschaftlichen wie im Praxiskontext verfügbaren Werkzeuge identifiziert. Anschließend werden diese Werkzeuge vergleichend analysiert, und entsprechend der Methode nach Nickerson et al. (2013) induktiv in eine umfassende Taxonomie überführt. Diese Taxonomie liefert einen Überblick über die differenzierenden Eigenschaften aktuell verfügbarer Software-Werkzeuge zur Canvas-basierten Geschäftsmodellentwicklung. Diese Taxonomie unterstützt Praktiker bei der Auswahl für sie geeigneter Werkzeuge und kann Forschern als Basis für weitere Forschung zur besseren Software-Unterstützung von Geschäftsmodell-Entwicklungsprozessen dienen.

## 2 Die Business Model Canvas

Die Business Model Canvas, auf dem die untersuchten Software-Werkzeuge aufsetzen, umfasst die folgenden neun Komponenten (Osterwalder und Pigneur 2010):

- *Wertangebot (Value Proposition)*: die Produkte und Dienstleistungen sowie die damit verbundenen Mehrwerte aus Sicht potenzieller Kunden (z. B. Auto, Strom)
- *Kundensegmente (Customer Segments)*: die Kunden, die mit dem Wertangebot adressiert werden sollen (z. B. Privatkunden, gewerbliche Kunden)
- *Kundenbeziehungen (Customer Relationships)*: die Art und Intensität der Beziehung zwischen Unternehmen und Kunden bei und nach dem Kauf (z. B. individuelle Beratung, Selbstbedienung)
- *Kanäle (Channels)*: die Vertriebs- und Distributionskanäle, über die die potenziellen Kunden erreicht werden sollen (z. B. stationärer Handel, Onlineplattform)

---

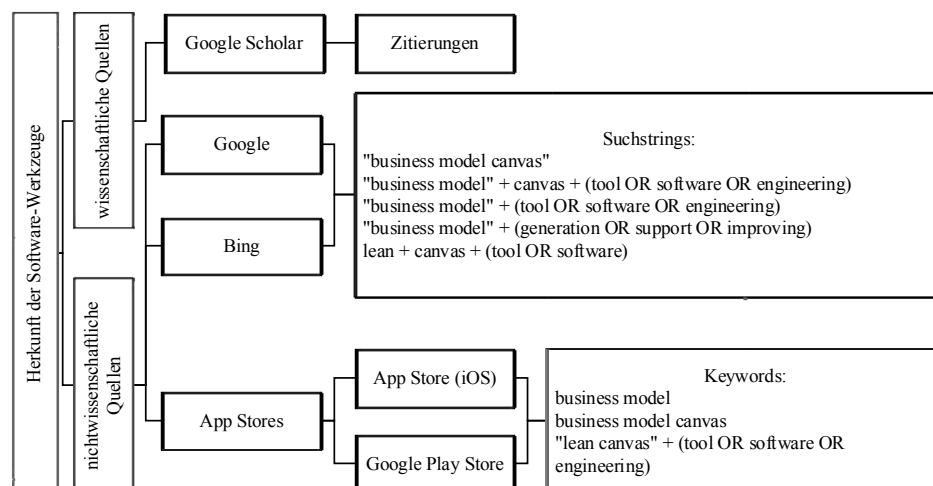
<sup>1</sup> Ermittelt am 03.10.2015.

- *Einnahmequellen (Revenue Streams)*: die Erlösquellen, über die das Unternehmen Umsatz generiert (z. B. Kaufpreis, Leasingrate)
- *Schlüsselressourcen (Key Resources)*: die zentralen Ressourcen, die für die Implementierung des Geschäftsmodells benötigt werden (z. B. Patente, spezielle Kompetenzen)
- *Schlüsselaktivitäten (Key Activities)*: die zentralen Aktivitäten (Kernprozesse), die im Geschäftsmodell durchgeführt werden müssen (z. B. Produktion, Marketing)
- *Schlüsselpartner*: die benötigten Partner, die vom Unternehmen nicht selbst durchgeführte Aktivitäten oder nicht vorhandene Ressourcen beisteuern (z. B. Lieferanten, Komplementäre)
- *Kostenstruktur (Cost Structure)*: die zentralen Kostenpositionen, die durch die Durchführung der Schlüsselaktivitäten, die Beschaffung von Schlüsselressourcen oder die Bezahlung von Schlüsselpartnern anfallen

### 3 Methodik

#### 3.1 Werkzeug-Identifikation

Ein Ziel der Arbeit ist es, einen möglichst umfassenden Überblick über derzeit existierende Werkzeuge zur Canvas-basierten Geschäftsmodellentwicklung zu liefern. Dieser Überblick sollte Werkzeuge aus der Wissenschaft und aus der Praxis umfassen – und dabei unabhängig von der durch das Werkzeug genutzten Software-Plattform sein. Der gewählte Suchprozess orientiert sich an der von Wolfswinkel et al. (2013) vorgeschlagenen Methode zur transparenten Literatursuche. Danach sind Datenquellen zu identifizieren sowie Auswahlkriterien und Suchstrings festzulegen, die im Laufe des Suchprozesses verfeinert werden können.



**Bild 1: Quellen der Werkzeugsuche**

Als Quellen für Werkzeuge aus der Wissenschaft wurde die wissenschaftliche Suchmaschine Google Scholar, für Werkzeuge aus der Praxis wurden die Internetsuchmaschinen Google und Bing, sowie die App Stores von Apple und Google genutzt. Auswahlkriterien waren, dass die gefundene Quelle ein Software-Werkzeug beschreibt mit dem auf Basis des Business Model Canvas Geschäftsmodelle entwickelt werden können, dass das Werkzeug vom Anbieter derzeit tatsächlich angeboten bzw. vertrieben wird (einige Quellen verwiesen auf Werkzeuge, die nicht mehr

zugreifbar sind), dass die Sprache der Benutzeroberfläche auf Deutsch oder Englisch ist und dass die Beschaffungskosten weniger als 30 € betragen oder eine kostenlose Testversion zur Verfügung gestellt wurde. Im Bild 1 werden die Quellen und die verwendeten Suchstrings aufgeführt.

Die Suche in Google Scholar mit trivialen Suchstrings wie "*business model canvas*" + (*tool OR software*) war erfolglos, da durch die Masse der Ergebnisse, bedingt durch die Suchbegriffe, die relevanten Veröffentlichungen nicht identifizierbar waren. Ausgehend von der Annahme, dass Arbeiten, die sich mit einem Software-Werkzeug zur Modellierung einer Business Modell Canvas befassen, die Dissertation von Alexander Osterwalder (2004) oder das Buch Business Model Generation (Osterwalder und Pigneur 2010) zitieren, konnte hier durch Rückwärtssuche ein Werkzeug identifiziert werden (Wolfswinkel et al. 2013). Um die Suche mit Google und Bing nicht durch Personalisierung der Suchergebnisse seitens der Suchmaschinen zu verfälschen, wurde jede Suche in Inkognito-Modus von Firefox durchgeführt. Die ersten 100 Treffer wurden untersucht. Hierbei gefundene Querverweise auf Webseiten wurden weiterverfolgt. Dabei wurden 15 Software-Werkzeuge identifiziert. In den App Stores von Apple und Google wurden 12 Apps identifiziert. Tabelle 1 listet alle identifizierten, aber nicht untersuchten Software-Werkzeuge auf. Insgesamt wurden abschließend 22 Werkzeuge untersucht.

Name des Software-Werkzeugs	Bemerkungen
Business Model	Android App, Fake-App, sammelt lediglich Emailadressen
Canvas de Modelo de Nego'cios	Android App, nicht in deutscher/ englischer Sprache verfügbar
Incuba for Startups	Nicht im deutschen iOS App Store verfügbar
Leanstack	Zum Zeitpunkt der Untersuchung nicht verfügbar
RunPat	Betrieb zum 22.08.2015 eingestellt
Strategyzer	Nicht finanzierbar (299\$/Projekt, keine Testversion)
TUZZit - Business Canvas (App)	Nicht im deutschen iOS App Store verfügbar

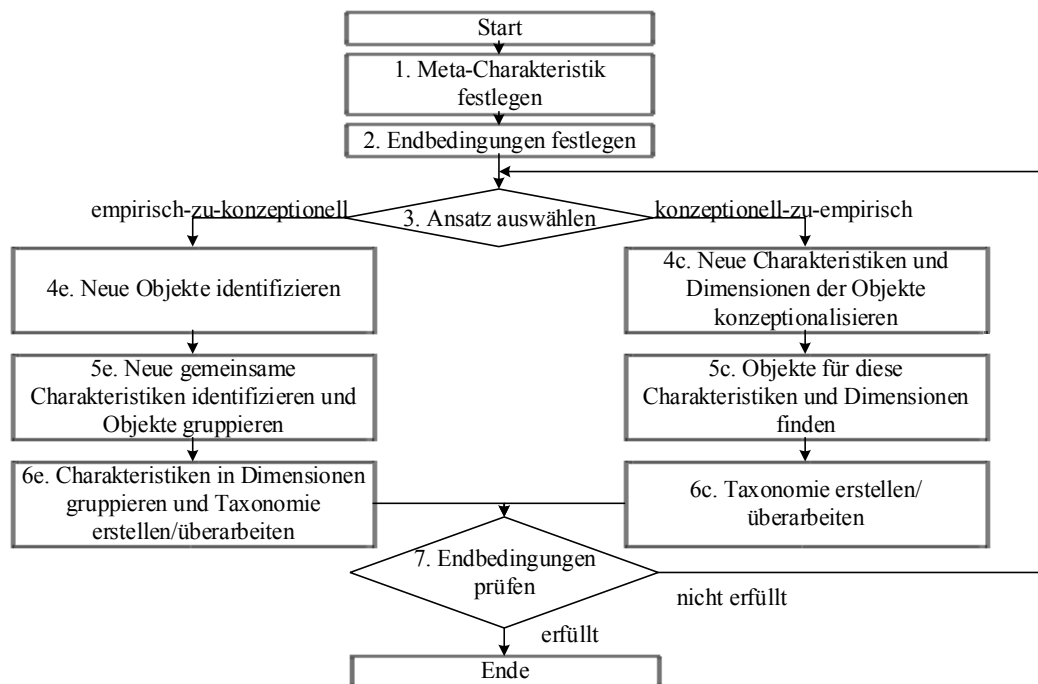
**Tabelle 1: Liste der identifizierten, jedoch nicht untersuchten Software-Werkzeuge**

### 3.2 Taxonomie-Entwicklung

Der folgende Abschnitt ist eine kurze Zusammenfassung der Methode zur Taxonomie-Entwicklung nach Nickerson et al. (2013). Unter Taxonomie versteht man eine Klassifizierung von Objekten. Objekte einer Domäne haben verschiedene Merkmale, hier Charakteristik genannt, die sich in Dimensionen gruppieren lassen. Eine Taxonomie lässt sich als Menge von Dimensionen, die wiederum Mengen von Charakteristiken sind, darstellen:  $T = \{D_i, i=1, \dots, n \mid D_i = \{C_{ij}, j=1, \dots, k_i; k_i \geq 2\}$  (Nickerson et al. 2013). Eine Taxonomie besteht aus mindestens einer Dimension und jede Dimension aus mindestens 2 Charakteristiken. Nach dem Prinzip des gegenseitigen Ausschlusses und der vollständigen Erschöpfung hat jedes Objekt genau eine Charakteristik in jeder Dimension (Nickerson et al. 2013). Die im Artikel vorgestellte systematische Methode zur Klassifizierung von Objekten unterstützt den Nutzer bei der Sicherstellung wichtiger Eigenschaften einer Taxonomie. Nach Nickerson et al. soll eine Taxonomie prägnant, robust, umfassend, erweiterbar und erklärend sein. Es soll also gewährleistet werden, dass so wenige Dimensionen und Charakteristiken wie möglich (*Prägnanz*) zum eindeutigen Differenzieren der Objekte (*Robustheit*) verwendet werden. *Umfassend* ist eine Taxonomie, wenn alle Objekte einer Domäne klassifiziert werden können und sie alle Dimensionen der zu untersuchenden Objekte beinhaltet. Falls neue Charakteristiken oder Dimensionen notwendig sind, z.B. durch Auftauchen eines neuen Objektes, kann sich die



Taxonomie erweitern lassen (*Erweiterbarkeit*). Diese Charakteristiken und Dimensionen sollen nicht jedes Detail der Objekte herausarbeiten, sondern Verständnis über alle Objekte schaffen (*Erklärend*). Das Ziel von Nickerson et al. (2013) war es, dass ihre Methode in angemessener Zeit und für spätere Taxonomie-Nutzer nachvollziehbar zu einer brauchbaren Taxonomie führt, und damit die Nutzung von ad hoc-Ansätzen zur Taxonomie-Erstellung reduziert. Außerdem soll die Methode leicht anzuwenden sein und verschiedene Ansätze bei der Taxonomie-Entwicklung in Betracht ziehen. Zu Beginn der Methode muss eine Meta-Charakteristik definiert werden (Schritt 1, siehe Bild 2). Diese Charakteristik soll die umfassendste Charakteristik sein aus der sich alle weiteren Charakteristiken logisch ableiten. Die Meta-Charakteristik richtet sich nach dem Verwendungszweck und den Nutzern der Taxonomie. Die von Nickerson et al. vorgestellte Methode ist iterativ und benötigt mindestens eine Endbedingung, um zu terminieren. Die Endbedingung muss vor der Durchführung festgelegt werden (Schritt 2). Hierbei wird zwischen subjektiven und objektiven Bedingungen unterschieden. Die subjektiven Endbedingungen sind die aufgezählten Qualitäten einer Taxonomie. Die objektiven Endbedingungen sind quantitativ messbar und terminieren auf diese Weise sicher.



**Bild 2: Methode zur Taxonomie-Entwicklung nach Nickerson et al. (2013)**

Nachdem die Meta-Charakteristik und die Endbedingungen festgelegt wurden, kann in den iterativen Erstellungsprozess gestartet werden. Vor jeder Iteration muss entschieden werden, ob ein empirisch-zu-konzeptionell-Ansatz oder ein konzeptionell-zu-empirisch-Ansatz gewählt wird (Schritt 3). Die Wahl des Ansatzes richtet sich nach Art der Objekte und Wissen des Forschers. Falls der Forscher über wenige Daten zu Objekten, aber viel Wissen über die Domäne verfügt, wird der konzeptionell-zu-empirisch-Ansatz empfohlen. In diesem Ansatz werden konzeptuell neue Charakteristiken und Dimensionen festgelegt (Schritt 4c) und dann anhand von Objekten empirisch überprüft (Schritt 5c). Abschließend wird die Taxonomie erstellt, erweitert oder überarbeitet (Schritt 6c). Hat der Forscher hingegen viele Daten über die Objekte und wenig Wissen über die Domäne, so wird ein empirisch-zu-konzeptionell-Ansatz empfohlen. Bei diesem Ansatz werden in

jeder Iteration neue Objekte identifiziert und der Taxonomie hinzugefügt (Schritt 4e). Dabei gilt es gemeinsame Charakteristiken zu identifizieren und zu gruppieren (Schritt 5e), um sie anschließend in (neue) Dimensionen einzuordnen und die Taxonomie zu erstellen, erweitern oder überarbeiten (Schritt 6e). Nach jeder Iteration der Ansätze werden die Endbedingungen überprüft (Schritt 7). Sind diese erfüllt, wird die Methode abgebrochen; sind sie es nicht, beginnt die nächste Iteration wieder mit der Auswahl des passenden Ansatzes (Schritt 3).

Nr.	Name	Leitfrage(n)
1	Prägnant	Ist die Taxonomie bedingt durch die Anzahl der Dimensionen aussagekräftig ohne schwerfällig oder überwältigend zu sein?
2	Robust	Erlauben es die Dimensionen und Charakteristiken Objekte eindeutig zu unterscheiden?
3	Umfassend	Können alle Objekte oder eine zufällige Stichprobe von Objekten aus der Domäne klassifiziert werden? Sind alle relevanten Dimensionen identifiziert?
4	Erweiterbar	Können neue Dimensionen oder eine Charakteristik einer bestehenden Dimension einfach hinzugefügt werden?
5	Erklärend	Was sagen die Dimensionen und Charakterisierten über ein Objekt aus?

**Tabelle 2: Subjektive Endbedingungen nach Nickerson et al. (2013)**

Nr.	Beschreibung
1	Alle Objekte oder eine repräsentative Stichprobe wurden untersucht
2	Kein Objekt wurde mit einem ähnlichen zusammengefasst oder in verschiedenen Objekte auf gesplittet
3	Jede Charakteristik wird mindestens von einem Objekt angenommen
4	In der letzten Iteration wurde keine Dimension oder Charakteristik hinzugefügt
5	In der letzten Iteration wurden keine Dimensionen oder Charakteristiken zusammengefasst oder aufgeteilt
6	Jede Dimension ist einzigartig und wird nicht wiederholt
7	Jede Charakteristik ist einzigartig und wird nicht wiederholt
8	Jede Kombination von Charakteristiken ist einzigartig und wird nicht wiederholt

**Tabelle 3: Objektive Endbedingungen nach Nickerson et al. (2013)**

### 3.3 Anwendung der Methode

Initial wurde die Meta-Charakteristik festgelegt. Entsprechend der Zielsetzung der Arbeit sind die Betrachtungsgegenstände der Taxonomie alle Software-Werkzeuge, die – basierend auf dem Geschäftsmodell-Verständnis des Business Model Canvas – die Entwicklung von Geschäftsmodellen unterstützen. Die Dimensionen der zu entwickelnden Taxonomie sollen Nutzern in der Praxis bei der Auswahl eines für sie geeigneten Werkzeugs helfen. Anschließend werden die Endbedingungen ausgewählt. Da alle Endbedingungen von Nickerson im gewählten Kontext sinnvoll anwendbar sind, wurde definiert, dass alle subjektiven und fast alle objektiven Endbedingungen zu erfüllen sind. Entgegen der von Nickerson et al. (2013) vorgestellten Methodik wurde das Prinzip des gegenseitigen Ausschluss der Charakteristiken ebenso wie die objektiven Endbedingungen 7 und 8 – die Einzigartigkeit der Charakteristiken und ihrer Kombinationen (siehe Tabelle 3) – zu Gunsten der Übersicht nicht strikt angewandt.

Darauf folgten fünf Iterationen mit einem empirisch-zu-konzeptionellen-Ansatz in dem die Werkzeuge nach Plattformen gruppiert mit Hilfe einer in Abschnitt 3.4 vorgestellten Fallstudie untersucht wurden. Um weitere Struktur innerhalb der Taxonomie zu schaffen, wurde noch eine Iteration mit einem konzeptionell-zu-empirische-Ansatz durchgeführt. Ziel dieser Iteration war es, die identifizierten Dimensionen auf einer höheren Ebene zu gruppieren, um die Übersichtlichkeit zu erhöhen. Hierfür wurde auf oberster Ebene angenommen, dass es Dimensionen gibt, die generelle Softwareeigenschaften beschreiben und solche, die Charakteristiken beinhalten, welche spezifisch für den Geschäftsmodellentwicklungsprozess relevant sind. Weiterhin wurde angenommen, dass der Geschäftsmodellentwicklungsprozess als kreativer Prozess betrachtet werden kann und damit der Phaseneinteilung kreativer Prozesse unterliegt. Um weitere Struktur zu schaffen, wurde daher das im Folgenden beschriebene Vierphasen-Modell kreativer Prozesse nach Howard (2008) genutzt, um die identifizierten, Geschäftsmodell-spezifischen Funktionen der identifizierten Werkzeuge zu gruppieren. Die erste Phase ist die Analysephase, sie dient der grundsätzlichen Klärung der Anforderungen einer Aufgabe. Während der (Ideen-) Generierungsphase sollen Ideen entwickelt, die dann in der Bewertungsphase bewertet und in der Implementierungsphase umgesetzt werden. Nach einer abschließenden konzeptionell-zu-empirischen Iteration terminierte die Methode, da keine neuen Objekte mehr zur Untersuchung vorlagen und sich konzeptuell keine Änderungen an der Taxonomie ergaben.

### 3.4 Werkzeug-Analyse

Jedes Werkzeug wurde einzeln untersucht. Um möglichst alle Aspekte der Software-Werkzeuge zu beleuchten, wurde eine Fallstudie entwickelt, in der – unterstützt durch die Software – ein Geschäftsmodell entwickelt werden soll. Die Idee war es, eine realitätsnahe Situation zu generieren. Die Fallstudie behandelt dabei ein Start-Up, in dem die Gründer an unterschiedlichen Orten gemeinsam arbeiten und sich mit externen Partnern über ihr Geschäftsmodell austauschen. In Abhängigkeit vom entsprechenden Werkzeug konnten die Anforderungen an verteiltes, gemeinsames Arbeiten und Entwicklung von neuen Ideen sowie der Kooperation mit externen Partner ausgetestet und dadurch neue Dimensionen oder Charakteristiken erschlossen werden.

## 4 Taxonomie

Im Folgenden sind zunächst die einzelnen Dimensionen der Taxonomie beschrieben. Anschließend ist die vollständige Taxonomie inklusive der Zuordnung der identifizierten Werkzeuge zu den Dimensionen der Taxonomie dargestellt.

### 4.1 Dimensionsbeschreibung

#### 4.1.1 Softwareeigenschaften

- *Plattform (Android, iOS, Windows, Linux/Mac, Browser-basiert)*: Benötigte Applikationsplattform der Software-Werkzeuge.
- *Kosten (kostenfrei, einmalig, regelmäßig)*: Häufigkeit der anfallenden Kosten.
- *Sprache (Deutsch, Englisch, Weitere)*: Verfügbare Sprachen innerhalb der Benutzeroberflächen.
- *Import/Export (Import, Export, nicht vorhanden)*: Eine Ausprägung in der Charakteristik Export bedeutet, dass Grafiken exportiert werden können. Existiert zusätzlich eine Ausprägung in

Import, können zusätzlich zu Grafiken auch Daten exportiert und diese Daten auch wieder importiert werden.

- *Bearbeitung (simultan, sequentiell)*: Simultanes Arbeiten ermöglicht es Teams gleichzeitig an einem Geschäftsmodell zu arbeiten. Bei sequentieller Arbeit ist, das nicht möglich und es muss bspw. das Endgerät weitergereicht werden.

#### 4.1.2 Geschäftsmodellentwicklungsprozess

##### (Ideen-)Generierung

- *Vorhandene Canvas-Varianten (BMC, Lean Canvas, Value Proposition, Weitere)*: Eine Übersicht über die angebotenen Canvas-Varianten. Die Business Model Canvas, sowie die Lean (Maurya 2012) und auch die Value Proposition Canvas (Osterwalder et al. 2015) sind bekannte Varianten. In der Charakteristik *Weitere* sind alle übrigen Canvas-Varianten abgebildet.
- *Benutzerunterstützung (Community, Wizard, Komponentenbibliothek, Erläuterungen, nicht vorhanden)*: In der Dimension Benutzerunterstützung sind Charakteristiken gelistet, die den Nutzer der Software-Werkzeuge bei der inhaltlichen Entwicklung seiner Geschäftsmodell-Ideen unterstützen. Während der Erstellung der Taxonomie wurden Erläuterungen der Bausteine, eine User Community und ein Wizard, der Schritt-für-Schritt durch den Erstellungsprozess eines Geschäftsmodells führt, identifiziert. Die Bausteinbibliothek erlaubt es, vordefinierte, domänenspezifische Bausteine auf der Canvas zu platzieren.
- *Darstellung (Post-It, Freitext)*: Unterscheidung zwischen strukturierter Darstellung mit Post-Its und Freitext in den Bausteinen. Post-Its lassen sich per Drag-and-Drop zwischen Bausteinen und innerhalb der Bausteine verschieben, dies ist bei Freitext nicht möglich.
- *Dokumentation (Annotation, Interview, Journal, nicht vorhanden)*: Annotation von Post-Its mit zusätzlichen Informationen. Archivierung von Interviews und festhalten von Entscheidungsprozessen und des Projektfortschritts mit einem Journal.
- *Kreativitätswerkzeuge (Ideensammlung, Hypothesen, nicht vorhanden)*: Mit Hypothesen ist eine Hypothesen-getriebene Geschäftsmodellentwicklung möglich. Entweder durch Aufstellen von separaten Hypothesen zum Geschäftsmodell oder mit Annotation an den Aussagen bzw. Post-Its. Die Charakteristik Ideensammlung beschreibt eine Funktionalität, die das schnelle Erstellen von Post-Its, die nicht sofort auf der Canvas platziert werden müssen, ermöglicht.
- *Versionierung (automatisch, manuell, nicht vorhanden)*: Versionierung erlaubt es, an bestimmte Zwischenstände des Geschäftsmodell-Erstellungsprozesses zurückzuspringen. Diese Punkte können manuell vom Nutzer oder automatisch von der Software gesetzt werden.

##### Bewertung

- *Bewertung (Finanzübersicht, Clusteranalyse, Umfragen, nicht vorhanden)*: Eine Finanzübersicht stellt mögliche Einnahmen den kalkulierten Kosten gegenüber. Eine Clusteranalyse kann in Kombination mit einer Komponentenbibliothek das Geschäftsmodell einem Marktsegment zuordnen. Mit Umfragen können innerhalb des Teams Ideen bewertet werden.

##### Implementierung

- *ArchiMate-Integration (vorhanden, nicht vorhanden)*: In der Phase der Implementierung bieten zwei Tools eine Integration mit dem ArchiMate-Standard (Iacob et al. 2012).

4.2 Taxonomie

Software-Werkzeug	Softwareeigenschaften					Phasen des Geschäftsmodellentwicklungsprozesses											
	Plattform	Kosten	Sprache	Import/Export	Bearbeitung	(Ideen-)Generierung										Bewertung	Implementierung
						Vorhandene Canvas Varianten	Benutzerunterstützung	Darstellung	Dokumentation	Kreativitätswerkzeuge	Versionierung	Bewertung	ArchiMate-Integration				
Android iOS Windows Linux/Mac Browser-basiert	kostenfrei einmalig regelmäßig	Weitere Englisch Deutsch	Import Export nicht vorhanden	simultan sequentiell	BMC Lean Canvas Value Proposition Weitere	Community Wizard Komponentenbibliothek Erläuterungen nicht vorhanden	Post-It Freitext	Annotation Interview Journal nicht vorhanden	Ideensammlung Hypothesen nicht vorhanden	automatisch manuell nicht vorhanden	Finanzübersicht Clusteranalyse Umfragen nicht vorhanden	vorhanden nicht vorhanden					
Business Model Canvas Startup	x	x	x	x	x	x		x	x		x	x		x	x		
Business Model Developer	x	x	x	x	x	x x	x x x	x	x		x	x	x x		x		
Canvas (BMC)	x	x	x	x	x	x	x	x	x		x	x		x	x		
BizCanvas	x	x	x	x	x	x x x x		x	x		x	x		x	x		
Business Canvas Model	x	x	x	x	x	x		x	x		x	x		x	x		
Business Model Canvas <sup>1</sup>	x	x	x	x	x	x	x	x	x		x	x		x	x		
Business Model Canvas <sup>2</sup>	x	x	x	x	x	x		x	x		x	x		x	x		
Business Model Toolbox	x	x	x	x	x	x x		x	x		x	x		x	x		
Canvas Model Design	x	x	x	x	x	x x x		x	x		x	x		x	x		
Archi	x x	x	x	x x	x	x		x	x		x	x		x	x		
Bizzdesign Enterprise Studio	x	x x	x x x	x x	x	x		x	x		x	x		x	x		
BMCANVAS		x x	x	x	x	x		x	x		x	x		x	x		
bmdesigner		x x	x	x	x	x		x	x		x	x		x	x		
Business Model Fiddle		x x	x x	x	x	x x		x	x		x x	x		x	x		
Canvanizer		x x	x	x	x	x		x	x		x x	x		x	x		
Canvanizer 2.0 (beta)		x x x	x	x	x	x		x	x		x x	x		x	x		
LeanLaunchLab		x	x	x	x	x		x x	x		x x x	x		x	x		
LeanMonitor		x	x	x	x	x x		x	x		x x x	x		x	x		
Lienzo		x	x x	x	x	x		x	x		x x	x		x	x		
Mural		x	x	x	x	x x x		x	x		x	x		x x	x		
TUZZit		x x x	x	x	x	x x		x	x		x	x		x	x		

Tabelle 4: Taxonomie zur Beschreibung und Auswahl von Business Model Canvas-basierten Software-Werkzeugen zur Geschäftsmodellentwicklung

[1] von Viacheslav Nechyporenko

[2] von Mathijs Vreeman

(Quellen zu den Software-Werkzeugen im Literaturverzeichnis)

## 5 Diskussion

Ziel der Arbeit war es, den Auswahlprozess von Software-Werkzeugen zur Geschäftsmodellentwicklung zu erleichtern. Während der Taxonomie-Entwicklung wurden die Dimensionen in eine allgemeine – generelle Software-Eigenschaften umfassende – Meta-Dimension und in vier weitere Meta-Dimensionen, die den kreativen Prozess der Geschäftsmodellentwicklung beschreiben, eingeordnet (Howard et al. 2008). Diese vier Meta-Dimensionen beschreiben mit ihren zugeordneten Dimensionen die Funktionen und Eigenschaften der Software-Werkzeuge.

In der Taxonomie wurden den Phasen (Ideen-)Generierung, Bewertung und Implementierung mindestens eine Dimension zugeordnet. Für die Analyse-Phase konnten keine Funktionen identifiziert werden. Die Phasen der Bewertung und Implementierung sind mit jeweils einer Dimension besetzt. Auffällig ist, dass in der Integration nur Desktop-Anwendungen (Windows, Linux/Mac) Funktionen in Form der Anbindung an den ArchiMate-Standard anbieten (Jacob et al. 2012) und so eine Enterprise Architecture Modelling ermöglichen. In der Phase der Bewertung bieten wenige Werkzeuge überhaupt Funktionen. Hier sticht der *Business Model Developer* hervor. Er bietet neben der auch in zwei anderen Werkzeugen vorhandenen Finanzübersicht zusätzlich eine Einteilung in Marktsegmente an (Clusteranalyse).

In der Phase der (Ideen-)Generierung bieten nur 6 von 22 Werkzeugen weiterführende Funktionalität. Insbesondere die Webbrowser-basierten Werkzeuge bieten in der Dimension der Kreativitätswerkzeuge vermehrt Unterstützung bei der Sammlung und Generierung neuer Ideen. Die Werkzeuge, die explizit den Lean-Ansatz (Maurya 2012) verfolgen, haben auch zusätzliche Funktionalität in der Dimension Dokumentation. Es können Interviews, Hypothesen und in einem Journal weitere Informationen den Projektfortschritt betreffend archiviert werden.

Betrachtet man die Dimension Bearbeitung aus der Meta-Charakteristik Softwareeigenschaften fällt auf, dass bei allen identifizierten mobilen Apps sowie einer Windows/Linux/Mac-Anwendung und einer Webanwendung nur sequentielles Arbeiten möglich ist. Für die Arbeit im Team sind diese Werkzeuge somit nur eingeschränkt nutzbar.

Es zeichnet sich ein Bild von zwei Gruppen von Software-Werkzeugen. Zum einen lassen sich die Individual- und Team-Werkzeuge anhand der Dimension Bearbeitung aufteilen. Zum anderen lassen sich die Software-Werkzeugen nach dem Grad ihrer Unterstützung beim Erstellen und (Weiter-)Entwickeln von Geschäftsmodellen unterteilen. Je mehr Ausprägungen (außerhalb von *nicht vorhanden*) in den Dimensionen Benutzerunterstützung und Kreativitätswerkzeuge vorhanden sind, desto stärker wird der Nutzer unterstützt.

## 6 Zusammenfassung und Ausblick

Die vorliegende Arbeit präsentiert eine Taxonomie zum Vergleich von Software-Werkzeugen mit denen Geschäftsmodelle mit Hilfe der Business Model Canvas entwickelt werden können. Auf Grundlage der Taxonomie konnten Gruppen von Software-Werkzeugen identifiziert und besondere Merkmale herausgearbeitet werden. Außerdem wurde untersucht welche Phasen kreativer Prozess die Software-Werkzeuge unterstützen. Der Auswahlprozess für Unternehmen kann durch den erstellten Überblick erleichtert werden. Eine genauere Aussage über das Ausmaß der Erleichterung lässt sich erst nach einer Evaluierung der Taxonomie in der Praxis treffen.

Die Taxonomie umfasst nur funktionale Eigenschaften von Software-Werkzeugen. Nicht-funktionale Eigenschaften sind schwer zu objektivieren. So könnte die Taxonomie nach

Eingrenzung von relevanten nicht-funktionalen Eigenschaften um entsprechende Dimensionen erweitert werden. Ein Ansatzpunkt für weitere Forschung könnte die Usability von Software in kreativen Prozessen sein.

## 7 Literatur

Archi, <http://www.archimatetool.com/>. Abgerufen am 24.8.2015

BizCanvas, iOS App Store. Abgerufen am 24.8.2015

BiZZdesign, <http://www.de.bizzdesign.com/tools/bizzdesign-architect/>. Abgerufen am 24.8.2015

BMCANVAS, <https://www.bmcanvas.com/>. Abgerufen am 24.8.2015

Bmdesigner, <http://bmdesigner.com/>. Abgerufen am 24.8.2015

Business Canvas Model, iOS App Store. Abgerufen am 24.8.2015

Business Model, iOS App Store. Abgerufen am 24.8.2015

Business Model Canvas, Mathijs Vreeman, iOS App Store. Abgerufen am 24.8.2015

Business Model Canvas, Viacheslav Nechyporenko, iOS App Store. Abgerufen am 24.8.2015

Business Model Canvas Startup, Google Play Store. Abgerufen am 24.8.2015

Business Model Developer, <http://uni-potsdam.de/dpm/bmapp.html>. Abgerufen am 24.8.2015

Business Model Fiddle, <https://www.bmfiddle.com/>. Abgerufen am 24.8.2015

Business Model Toolbox, iOS App Store. Abgerufen am 24.8.2015

Canvanizer/ Canvanizer 2.0 (beta), <https://canvanizer.com/>. Abgerufen am 24.8.2015

Canvas (BMC), Google Play Store. Abgerufen am 24.8.2015

Canvas de Modelo de Nego'cios, Google Play Store. Abgerufen am 24.8.2015

Canvas Model Design, iOS App Store. Abgerufen am 24.8.2015

GE (2014) GE Global Innovation Barometer 2014, General Electric

Howard TJ, Culley SJ, Dekoninck E (2008) Describing the creative design process by the integration of engineering design and cognitive psychology literature. *Design studies*, 29(2):160-180

Iacob M, Jonkers H, Lankhorst M, Proper E, Quartel DAC (2012) ArchiMate 2.0 Specification

IBM (2008) The enterprise of the future: Global CEO study, IBM Institute for Business Value

Incuba for Startups, iOS App Store. Abgerufen am 24.8.2015

LeanLaunchLab, <https://www.leanlaunchlab.com/>. Abgerufen am 24.8.2015

LeanMonitor, <http://leanmonitor.com/>. Abgerufen am 24.8.2015

Leanstack, <http://leanstack.com/>. Abgerufen am 24.8.2015

Lienzo, <http://lienzo.biz/>. Abgerufen am 24.8.2015

Mural, <https://mural.ly/>. Abgerufen am 24.8.2015

- Nickerson RC, Varshney U, Muntermann J (2013) A method for taxonomy development and its application in information systems. *European Journal of Information Systems*, 22(3):336-359
- Osterwalder A (2004) The business model ontology: A proposition in a design science approach. Dissertation
- Osterwalder A, Pigneur Y (2010) *Business model generation: a handbook for visionaries, game changers, and challengers*. Wiley, Hoboken, NJ
- Osterwalder A, Pigneur, Y (2013) Designing business models and similar strategic objects: The contribution of IS. *Journal of the Association for Information Systems*, 14(Special Issue):237-244
- Osterwalder A, Pigneur Y, Bernarda G, Smith A (2015) *Value Proposition Design: How to Create Products and Services Customers Want*. Wiley, Hoboken, NJ
- Maurya A (2012) *Running lean: iterate from plan A to a plan that works*. O'Reilly Media, Inc., Sebastopol, CA
- RunPat, <https://www.runpat.com/>. Abgerufen am 24.8.2015
- Shneiderman B (2007) Creativity support tools: Accelerating discovery and innovation. *Communications of the ACM*, 50(12):20-32
- Strategyzer (2015) *The business model canvas: Why and how organizations around the world adopt it*, Strategyzer & Business Models Inc.
- Strategyzer, <https://strategyzer.com/app>. Abgerufen am 24.8.2015
- Teece DJ (2010) Business models, business strategy and innovation. *Long Range Planning*, 43(2-3):172-196
- TUZZit, <https://www.tuzzit.com/>. Abgerufen am 24.8.2015
- TUZZit - Business Canvas, iOS App Store. Abgerufen am 24.8.2015
- Wolfswinkel JF, Furtmueller E, Wilderom CP (2013) Using grounded theory as a method for rigorously reviewing literature. *European Journal of Information Systems*, 22(1):45-55
- Zott C, Amit R, Massa L (2011) The business model: recent developments and future research. *Journal of management*, 37(4):1019-1042



# Engineering-Prozess zur Gestaltung eines CPS für Instandhalter

Michael Wächter<sup>1</sup> und Angelika C. Bullinger<sup>1</sup>

<sup>1</sup> Technische Universität Chemnitz, Professur Arbeitswissenschaft und Innovationsmanagement, michael.waechter@mb.tu-chemnitz.de

## Abstract

Im Zuge der vierten industriellen Revolution entstehen durch die Verbindung von Informations- oder Cybersystemen mit physischen Systemen neue Möglichkeiten der Mensch-Maschine-Interaktion. Die Gestaltung von gebrauchstauglichen Hardware-Interfaces, die in der Anwendung hohe Akzeptanz finden, stellt dabei eine besondere Herausforderung dar. Während die Literatur zahlreiche Vorgehensmodelle zur Gestaltung gebrauchstauglicher Softwareoberflächen bietet, fehlen Hinweise für die Gestaltung gebrauchstauglicher Hardware-Elemente für die Benutzerschnittstelle. Aufbauend auf dem Usability-Engineering und Design Science Research schlägt dieser Beitrag einen iterativen Engineering-Prozess zur Gestaltung gebrauchstauglicher Hardware-Interfaces am Beispiel eines mobilen Assistenzsystems für Instandhalter vor. Dabei werden die einzelnen Phasen des Prozesses erläutert und methodisch untersetzt.

## 1 Einleitung

Die vierte industrielle Revolution steht für die intelligente Vernetzung von Maschinen und Anlagen entlang der gesamten Wertschöpfungskette. In der Vision stimmen cyber-physische Systeme (CPS) durch die permanente Erfassung von Zustandsdaten ihre Aufträge, Ressourcen und relevante Informationen autonom untereinander ab und koordinieren eigenständig notwendige Inspektionen, Wartungen und Instandhaltungsmaßnahmen (Bauer et al. 2014). Kontinuierlich verfügbare Informationen sollen dafür anwenderspezifisch aufbereitet und für die Optimierung des Produktionsprozesses genutzt werden (Botthof und Hartmann 2015). Der Mensch, der z.B. als Instandhalter die Verfügbarkeit der kapitalintensiven, smarten und vernetzten Subsysteme der CPS verantwortet, wird dabei noch stärker zum Erfolgsfaktor (Geisberger und Broy 2012). Das Szenario zukünftiger, effizienter Instandhaltung sieht dazu vor, dass alle relevanten Zustandsinformationen, Reparaturanleitungen und Dokumentationen mobil abrufbar zur Verfügung stehen (Scheer 2013). In der Gegenwart erfolgt dieser Informationsabruf häufig papierbasiert bzw. telefonisch; mobile Endgeräte sind in der Produktion noch selten im Einsatz. Deren Einsatzmöglichkeiten steigen in Folge der großen Datenverfügbarkeit jedoch zunehmend (Ganschar et al. 2013), so dass Bedarf an einem Engineering-Prozess für ihre Gestaltung besteht (Botthof und Hartmann 2015). Dieser kann aktuell jedoch weder aus Wissenschaft noch Praxis beantwortet werden (Gerhard 2009), was

überrascht, da im Bereich der Softwaregestaltung eine Vielzahl an Richtlinien und Leitfäden existieren. Aufgrund der identifizierten Lücke wird im vorliegenden Beitrag der Vorschlag eines Engineering-Prozesses zur Gestaltung von CPS am Beispiel eines mobilen Assistenzsystems für Instandhalter präsentiert.

Dieser research-in-progress Beitrag ist im Folgenden gegliedert in eine kurze Darstellung der methodischen Vorgehensweise zur Entwicklung des Engineering-Prozesses und dessen Präsentation mit den fünf grundlegenden Schritten. Die Ergebnisse der ersten drei Prozessschritte werden präsentiert und die Prozessbausteine Lösungskonstruktion und Evaluation methodisch unterfüttert. Ein Ausblick auf nächste Forschungsschritte schließt den Beitrag.

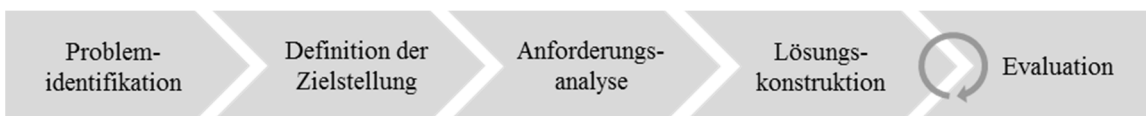
## 2 Methodische Vorgehensweise

Das verwendete Forschungsdesign orientiert sich am Forschungsparadigma der Design Science Research (DSR) nach Hevner et al. (2004). Fehlende Leitfäden zur Gestaltung und Evaluation von Artefakten wie z.B. einem mobilen Endgerät für die Produktion gelten als wissenschaftliche Lücke innerhalb der DSR-Community (Gacenga et al. 2012). Insbesondere zur Evaluation von Artefakten, d.h. der strategischen und methodischen Vorgehensweise, herrscht Uneinigkeit (Carlsson et al. 2011). Bewertungsrichtlinien sind aber notwendig, um die wissenschaftliche Stringenz von DSR kontinuierlich zu verbessern (Winter 2008).

Auf Basis einer Literaturanalyse zu DSR und Hardware-Usability wird deshalb ein Engineering-Prozess entwickelt, der zur Artefakt-Gestaltung und –Evaluation Prozessschritte aus der Literatur nutzt und Methoden aus DSR und dem Usability-Engineering (Hevner et al. 2004; Vaishnavi und Kuechler 2015; Nielsen 1993) integriert. Gemäß dem Design Science Paradigma soll durch diese Vorgehensweise die Lösung der praxisrelevanten Problemstellung (gebrauchstaugliches mobiles Assistenzsystem für Instandhalter) erreicht und die akademische Wissensbasis in Form eines Engineering-Prozesses für CPS erweitert werden.

## 3 Engineering-Prozess zur Gestaltung von CPS

In der Analyse von 14 DSR-Prozessen konnten die fünf Kernelemente Problemidentifikation, Definition der Zielstellung, Anforderungsanalyse, Lösungskonstruktion und Evaluation identifiziert werden. Die Phasen der *Lösungskonstruktion* und *Evaluation* erfolgen iterativ (s.u.).



**Bild 1: Elemente des Engineering-Prozesses**

Die Phasen *Problemidentifikation*, *Definition der Zielstellung* wurden im Rahmen des Forschungsprozesses bereits bearbeitet und aufbauend auf den Ergebnissen wurde eine *Anforderungsanalyse* durchgeführt. Die Ergebnisse werden im Folgenden erläutert. Die iterative *Lösungskonstruktion* und *Evaluation* stellen Schlüsseldisziplinen des DSR dar (Nunamaker und Chen 1990; March und Smith 1995; Hevner et al. 2004; Peffers et al. 2007) und werden daher nachfolgend besonders ausführlich betrachtet.

### 3.1 Problemidentifikation und Zielstellung

Ein Großteil der analysierten Literatur zum Usability-Engineering widmet sich der Gestaltung von Softwareanwendungen und Bildschirmarbeit (Sarodnick und Brau 2006). Dabei finden Mensch-Maschine-Interaktionen auch im tangiblen Bereich statt und weisen speziell in der Produktion besondere Anforderungen auf. Parallel ist zu beobachten, dass mobile Endgeräte, ursprünglich für den privaten Gebrauch entwickelt, Einzug in den Produktionsbereich halten (Ganschar et al. 2013). Aktuell existieren jedoch keine Hinweise zur Gestaltung ihrer Hardware-Schnittstelle zum Mensch als Nutzer. Es ist daher das Ziel, einen Engineering-Prozess zu entwickeln, anzuwenden und zu evaluieren, mit dem gebrauchstaugliche Hardware-Interfaces gestaltet werden können.

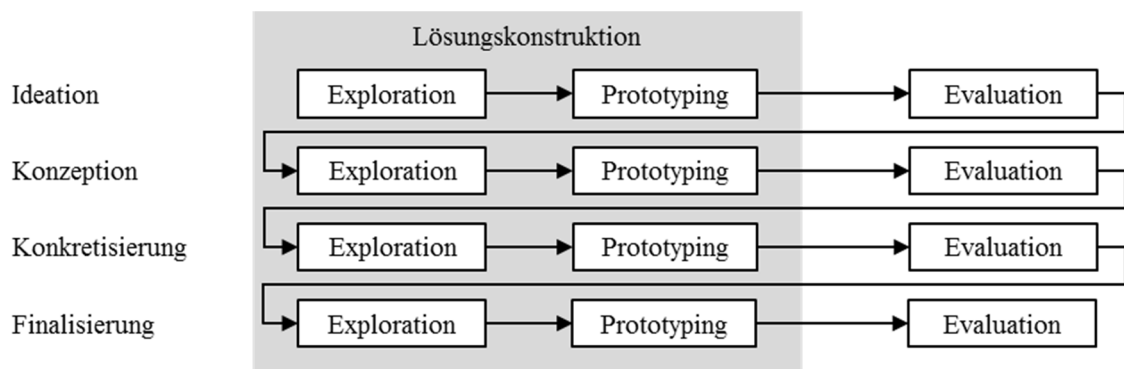
### 3.2 Anforderungsanalyse

Aufbauend auf der Zielstellung wurde eine Anforderungs- und Nutzerkontextanalyse in Unternehmen der Automobilindustrie, der Automobil-Zulieferindustrie sowie der Windkraftbranche durchgeführt. Im Rahmen von Expertenbefragungen, Prozessbeobachtungen und Dokumentenanalysen konnten die Anforderungen an das Hardware-Interface durch 18 leitfadengestützte Interviews mit Instandhaltern erfasst werden. Bei Verfügbarkeit verschiedener mobiler Endgeräte befürworteten die Befragten den Einsatz eines Tablet-artigen Assistenzsystems mit einer Displaydiagonale von acht bis zehn Zoll. Nichtfunktionale Anforderungen umfassen eine robuste Bauweise für den Einsatz im industriellen Umfeld und eine intuitive Bedienung. Im Zuge der Beobachtungen konnten zudem folgende funktionale Anforderungen identifiziert werden: Transportmöglichkeit; Bedienung mittels Touchscreen und physischer Bedienelemente; Stellbarkeit auf ebenen Flächen sowie Anheftbarkeit an Maschinen und Anlagen während des Instandhaltungsprozesses (Wächter und Bullinger 2015).

Diese Ergebnisse bilden den Anforderungskatalog für das Hardware-Interface des CPS und stellen die Basis für die iterative Lösungskonstruktion und Evaluation dar.

## 4 Iterative Lösungskonstruktion und Evaluation

Die *Lösungskonstruktion* (mit den beiden Teilschritten Exploration und Prototyping) und *Evaluation* erfolgen in den vier Iterationen *Ideation*, *Konzeption*, *Konkretisierung* und *Finalisierung*. Dabei werden pro Iteration Gestaltung und Evaluation unterschiedlich angelegt (Bild 2). Die dazu notwendigen Grundlagen, Teilnehmer, eingesetzte Methoden und Verfahren sowie die Ergebnisse werden im Folgenden näher erläutert.



**Bild 2: Iterative Lösungskonstruktion und Evaluation**

#### 4.1 Iteration 1: Ideation

In der ersten Iteration werden zunächst Ideen zur Umsetzung der erhobenen Anforderungen gesammelt (Tabelle 1). Um eine möglichst hohe Ideenanzahl zu generieren und eine bereits vorhandene gedankliche Ausrichtung oder Fixierung zu vermeiden, sollten in dieser Phase auch Nicht-Anwender teilnehmen (Wächter und Bullinger 2015).

Phase 1: Ideation	Lösungskonstruktion		Evaluation
	Exploration	Prototyping	
Grundlage	Anforderungskatalog	Gestaltungsideen	Gestaltungsentwürfe
Teilnehmer	Nicht-Anwender		Experten der Produktgestaltung
Methoden, Wissen, Verfahren	Fokusgruppe, Brainstorming	Fokusgruppe, Design Thinking	Workshop, Heuristik
Ergebnis	Gestaltungsideen	Gestaltungsentwürfe	kombinierter Gestaltungsentwurf

**Tabelle 1: Lösungskonstruktion und Evaluation Iteration 1: Ideation**

Die *Exploration* für die Umsetzung der erhobenen funktionalen Anforderungen erfolgt mittels Brainstorming. Anschließend erhalten die Teilnehmer die Möglichkeit, ihre Ideen unter Verwendung von Modelliermasse *prototypisch* umzusetzen. Die modellierten Gestaltungsentwürfe dienen als Grundlage zur qualitativen Bewertung und Kategorisierung der einzelnen Funktionen durch Experten zur Produktergonomie. Die heuristische *Evaluation* durch Experten der Produktgestaltung weist eine geeignete Methode zur Evaluation der entstandenen Gestaltungsentwürfe am Ende der ersten Iteration auf (Nielsen 1993). Die gewonnenen Ergebnisse in Form eines kombinierten Gestaltungsentwurfs bilden die Grundlage für die Entwicklungsphase der Konzeption.

#### 4.2 Iteration 2: Konzeption

In der zweiten Iterationsphase werden für die *Lösungskonstruktion* zunächst verwandte Produktlösungen analysiert. Diese erweitern, zusammen mit den Erkenntnissen zur Gestaltung von Stellteilen (DIN EN 894-3), die Grundlage für weitere Konzeptentwürfe zur Produktform (Tabelle 2). Der kombinierte Gestaltungsentwurf aus der ersten Iteration und die Konzeptentwürfe werden dazu mittels CAD konstruiert.

Phase 2: Konzeption	Lösungskonstruktion		Evaluation
	Exploration	Prototyping	
Grundlage	Kombinierter Gestaltungsentwurf	Verschiedene Konzeptentwürfe	physische Konzeptmodelle
Teilnehmer	Experten		Anwender
Methoden, Wissen, Verfahren	Fokusgruppe, DIN EN 894-3, verwandte Produktarten	CAD, 3D-Druck digit. Menschmodelle, DIN 33402-2	Fokusgruppe, Fragebogen (CQH), Lautes Denken
Ergebnis	verschiedene Konzeptentwürfe	physische Konzeptmodelle	abgestimmtes Konzeptmodell

**Tabelle 2: Lösungskonstruktion und Evaluation Iteration 2: Konzeption**

Unter Verwendung der Körpermaße des Menschen (DIN 33402-2) erfolgt zudem die Anpassung hinsichtlich anthropometrischer Variablen. Um die Anwender schon frühzeitig in den Entwicklungsprozess einbeziehen zu können, eignet sich die Herstellung von hardwaretechnischen *Prototypen*. So können schon während des Entwicklungsprozesses die Gestaltung, Haptik und Anforderungserfüllung getestet und *evaluiert* werden. Mehrere Konzeptmodelle erlauben darüber hinaus einen Vergleich. Mit Hilfe des Comfort Questionnaire of Hand-Tools (Kuijt-Evers et al. 2004) können die Meinungen der Anwender zur Qualität der Konzeptmodelle aufgenommen und anschließend ausgewertet werden. Ergebnis dieser Phase ist die Festlegung einer Produktform als Grundlage für die konkrete Ausgestaltung der weiteren Funktionen aus der Anforderungsanalyse.

#### 4.3 Iteration 3: Konkretisierung

Die dritte Iteration konkretisiert im Rahmen der *Lösungskonstruktion* das abgestimmte Konzeptmodell und erweitert dieses um weitere Funktionen, wie physische Bedienelemente, aus der Anforderungsanalyse. Hier können im Teilschritt der Exploration mit Hilfe des morphologischen Kastens verschiedene Anordnungen und Ausprägungen der Bedienelemente entstehen, die anschließend im Rahmen der Prototypenerstellung im CAD mit Hilfe von digitalen Menschmodellen detailliert werden (Tabelle 3).

Phase 3: Konkretisierung	Lösungskonstruktion		Evaluation
	Exploration	Prototyping	
Grundlage	abgestimmtes Konzeptmodell	verschiedene Konzeptmodelle	Geometrieprototypen
Teilnehmer	Experten		Anwender
Methoden, Wissen, Verfahren	Fokusgruppe, Morphologischer Kasten	CAD, 3D-Druck digit. Menschmodelle,	Fragebögen (SUS, CQH), Lautes Denken
Ergebnis	verschiedene Konzeptmodelle	Geometrieprototypen	abgestimmter Geometrieprototyp

**Tabelle 3: Lösungskonstruktion und Evaluation Iteration 3: Konkretisierung**

Dabei werden die Greifräume und Greifarten nach DIN 33402-2 (2005) berücksichtigt und in einen primären, sekundären und tertiären Greifraum eingeteilt, um die Gestaltung und Anordnung der physischen Bedienelemente vorzunehmen. Mittels 3D-Drucktechnologien werden anschließend die *Geometrieprototypen* gefertigt und methodisch analog dem Teilschritt Konzeption evaluiert. Zusätzlich erfolgt eine Bewertung mit Hilfe der System Usability Scale (Brooke 1996), um im Ergebnis einen abgestimmten Geometrieprototypen zu erreichen.

#### 4.4 Iteration 4: Finalisierung

In der letzten Iterationsschleife erfolgt die Finalisierung des abgestimmten Geometrieprototyps zu einem Funktionsprototyp (Tabelle 4). Dazu werden im Rahmen der *Lösungskonstruktion* letzte Anpassungen vorgenommen und ein finaler Prototyp gefertigt. Als Grundlage dienen die Evaluationsergebnisse der dritten Iteration. Der *Funktionsprototyp* repräsentiert einen Großteil der Anforderungen und dient als Grundlage für die Evaluation des CPS im späteren Anwendungsbereich. Zum Beispiel ist nun das Bedienelement inklusive einer Testanwendung umgesetzt.

Phase 4: Finalisierung	Lösungskonstruktion		Evaluation
	Exploration	Prototyping	
Grundlage	abgestimmter Geometrieprototyp	angepasster Geometrieprototyp	Funktionsprototyp
Teilnehmer	Experten		Anwender
Methoden, Wissen, Verfahren	Fokusgruppe	CAD, 3D-Druck	Fragebögen (SUS, CQH), Lautes Denken, Usability-Test
Ergebnis	angepasster Geometrieprototyp	Funktionsprototyp	evaluierter Funktionsprototyp

**Tabelle 4: Lösungskonstruktion und Evaluation Iteration 4: Finalisierung**

Mit Hilfe des entwickelten Funktionsprototyps erfolgt die *Evaluation* des CPS anhand der aus den vorherigen Iterationen bekannten Evaluationsmethoden. Der wiederkehrende Einsatz der eingesetzten Evaluationsmethoden ermöglicht dabei einen Vergleich der verschiedenen Entwicklungsstadien. Der evaluierte Funktionsprototyp ist das Ergebnis der letzten Phase des Engineering-Prozesses.

## 5 Zusammenfassung und Ausblick

Im vorliegenden Beitrag wird ein Engineering-Prozess zur Gestaltung gebrauchstauglicher CPS vorgestellt. Am Beispiel der Gestaltung eines mobilen Assistenzsystems für Instandhalter werden fünf Prozessschritte dargestellt. Die Problemidentifikation und die daraus abgeleitete Zielstellung bilden dabei den Ausgangspunkt für eine zielgerichtete Anforderungsanalyse. Die bereits erhobenen Anforderungen der Anwender werden in den beiden Phasen des Engineering-Prozesses im Rahmen der Lösungskonstruktion und Evaluation iterativ umgesetzt. Die methodische Vorgehensweise während der Entwicklung und die Einbeziehung der Anwender in die Evaluation gewährleistet die Integration von Anwenderwissen in die jeweilig folgende Iteration von Anfang an. So wird auch die spätere Akzeptanz des CPS abgesichert.

Aktuell befindet sich der hier präsentierte research-in-progress zur Gestaltung eines mobilen Assistenzsystems für Instandhalter in dem Teilschritt Konzeption, d.h. der zweiten Iteration von Lösungskonstruktion und Evaluation. Die finale Gestaltung des mobilen Endgeräts und die daran anschließende Evaluation des Engineering-Prozesses als solches stehen noch aus.

*Danksagung:*

Dieser Beitrag entstand im Rahmen des Projekts „Ressourcen-Cockpit für Sozio-Cyber-Physische Systeme“. Dieses Forschungs- und Entwicklungsprojekt wird mit Mitteln des Bundesministeriums für Bildung und Forschung (BMBF) im Rahmenkonzept „Forschung für die Produktion von morgen“ gefördert und vom Projektträger Karlsruhe (PTKA) betreut. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autoren.

## 6 Literatur

Bauer W, Schlund S, Marrenbach D (2014) Industrie 4.0 – Volkswirtschaftliches Potenzial für Deutschland: [Studie]. <https://www.bitkom.org/Bitkom/Publikationen/Industrie-40.html>. Abgerufen am 18 September 2015

- Botthof A, Hartmann EA (2015) *Zukunft der Arbeit in Industrie 4.0*. Springer Berlin Heidelberg, Berlin, Heidelberg
- Brooke J (1996) SUS: a 'quick and dirty' usability scale. In: Jordan PW (Hrsg.) *Usability evaluation in industry*. Taylor & Francis, London, Bristol, Pa., S. 189–194
- Carlsson SA, Henningsson S, Hrastinski S, Keller C (2011) Socio-technical IS design science research: developing design theory for IS integration management. *Information Systems and e-Business Management* 9(1):109–131
- DIN 33402-2 (2005) *Ergonomie - Körpermaße des Menschen*
- DIN EN 894-3 (2010) *Ergonomische Anforderungen an die Gestaltung von Stellteilen*
- Gacenga F, Cater-Steel A, Toleman M, Tan W (2012) A Proposal and Evaluation of a Design Method in Design Science Research. *Electronic Journal of Business Research Methods* 10(2):89–100
- Ganschar O, Gerlach S, Hämmerle Moritz, Krause T, Schlund S (2013) *Produktionsarbeit der Zukunft - Industrie 4.0: [Studie]*. Fraunhofer Verlag, Stuttgart
- Geisberger E, Broy M (2012) *agendaCPS: [Studie]*, vol 1. Springer, Berlin, Heidelberg
- Gerhard M (2009) Akzeptanz mobiler räumlicher Assistenzsysteme. In: Großmann U (Hrsg.) *Innovative mobile Technologien und Anwendungen*. LIT, Berlin, S. 39–50
- Hevner AR, March ST, Park J, Ram S (2004) Design Science in Information Systems Research. *MIS Quarterly* 28(1):75–105
- Kuijt-Evers LF, Groenesteijn L, de Looze, Michiel P., Vink P (2004) Identifying factors of comfort in using hand tools. *Applied ergonomics* 35(5):453–458
- March ST, Smith GF (1995) Design and natural science research on information technology. *Decision Support Systems* 15(4):251–266
- Nielsen J (1993) *Usability Engineering*. Morgan Kaufmann Publishers, San Francisco, Calif.
- Nunamaker JF, Chen M (1990) Systems Development in Information Systems Research. In: 23rd Hawaii International Conference on System Sciences (HICSS): Proceedings. IEEE, S. 631–640
- Peffers K, Tuunanen T, Rothenberger MA, Chatterjee S (2007) A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems* 24(3):45–77
- Sarodnick F, Brau H (2006) *Methoden der Usability Evaluation: Wissenschaftliche Grundlagen und praktische Anwendung*, 1. Aufl. *Praxis der Arbeits- und Organisationspsychologie*. Huber, Bern
- Scheer A (2013) *Industrie 4.0: Wie sehen Produktionsprozesse im Jahr 2020 aus?* IMC AG
- Vaishnavi V, Kuechler W (2015) *Design Science Research Methods and Patterns: Innovating information and communication technology*, Second edition. CRC Press
- Wächter M, Bullinger AC (2015) Gestaltung gebrauchstauglicher Assistenzsysteme für Industrie 4.0. In: Weisbecker A, Burmester M, Schmidt A (Hrsg.) *Mensch und Computer 2015: Workshopband*, Berlin, S. 165–169
- Winter R (2008) Design Science Research in Europe. *European Journal of Information Systems* 17(5):470–475





# The Impact of Anonymity on Persuasiveness in Online Participation

Thomas Wagenknecht<sup>1</sup>, Timm Teubner<sup>2</sup>, and Christof Weinhardt<sup>2</sup>

<sup>1</sup> FZI Forschungszentrum Informatik, Berlin, wagenknecht@fzi.de

<sup>2</sup> Karlsruher Institut für Technologie, Institut für Informationswirtschaft und Marketing, timm.teubner@kit.edu

## Abstract

Online participation is a powerful means of involving citizens, members, and employees in sovereign, organizational, and corporate discussion and decision processes. However, it immanently poses the question how to deal with user privacy. Anonymity, on the one hand, offers user protection against repression and can free reticent members. On the other hand, however, it faces the challenge of decreased levels of social presence and credibility as well as the risk of negative dis-inhibition (foul language, hoax). Design alternatives range from full anonymity to full user identification. In this article, we propose a study design to investigate the effect of user anonymity on communication persuasiveness, based on a set of fictional discussion threads with two opposing parties arguing in favor of their opinion. We systematically vary the degree to which users of both sides appear anonymous or identifiable. Our conjectures are rooted in the communication literature and build on the concepts of social presence and user credibility, representing highly relevant design issues for operators of online discussion and decision processes.

## 1 Introduction

Small and large firms alike are increasingly operating in new ways to adapt to the changing needs and demands of their staff. As today's employees ask for more independence in terms of how, when, and where they work, computer supported cooperative work (CSCW) becomes increasingly important (Behrendt et al., 2014). Since the 1990s, companies are using group decision support systems (GDSS) to enable decision-making on diverse topics for spatially and timely dispersed teams (Postmes and Lea, 2000). In GDSS, employers sometimes introduce anonymity in the decision-making process (Rains, 2007). It protects user privacy and reduces several negative effects that usually occur in group decision-making. For instance, it can lead to reduced conformity as well as decreased ownership biases (Sia et al., 2002).

On one hand, lower-level, yet knowledgeable employees might be reluctant to argue against their superior managers in a debate because they fear repression. This might lead to fewer arguments expressed, suboptimal decisions, and hence, low employee satisfaction. Thus, an option for anonymity in GDSS might be useful. On the other hand though, anonymity also yields new

challenges to the facilitators of GDSS. For instance, anonymous discussions were found to be more polarizing (Sia et al., 2002) and to include more hate speech and foul language (Sæbø et al., 2010).

Beyond that, the users' levels of anonymity may alter the way their arguments and ideas are perceived by others in terms of persuasiveness. Thus, it is of high interest for GDSS facilitators to learn to what extent user anonymity is perceived as engaging, entertaining, and valuable. There is, to the best of our knowledge, no systematic attempt to describe the impact of user anonymity on communication persuasiveness in online participation and debates.

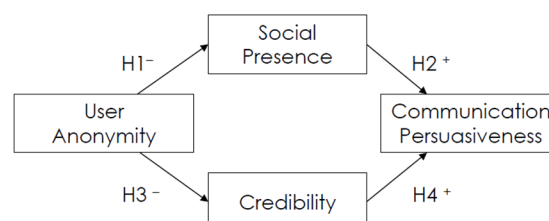
In this research in progress article, we hence outline a path to investigate the effect of user anonymity on communication persuasiveness. To this end, we propose an online survey, employing a set of fictional discussions in which two opposing parties argue in favor of their respective opinion. We systematically vary the degree to which users of both sides are represented, either anonymous or identifiable by their photograph. In doing so, this study makes one main contribution to the information systems literature, as it presents a rigorous approach to investigating user anonymity in computer-mediated communication (CMC), locating this research within the rich literature on (computer mediated) communication theory and persuasive communication.

The remainder of this article is organized as follows. In Section 2, we illustrate the theoretical background and, based on theories of social presence and de-individualization, derive our research model and hypotheses. Section 3 presents our study outline, including treatment design and stimuli material. We discuss potential issues and ramifications in Section 4.

## 2 Theoretical Background and Related Work

Anonymity in GDSS has been investigated extensively in the past decades. Scholars claim that anonymity influences decision quality, quantity of ideas and solutions as well as user satisfaction (Postmes and Lea, 2000). Therefore, anonymity might cause both positive as well as negative effects. For instance, by removing information on age, sex, and other demographic factors, anonymity can improve the quality of a debate (Sæbø et al., 2010). Anonymous groups have also been found to propose more novel ideas to a problem (Sia et al., 2002). While these are highly relevant results from a GDSS operator's perspective, individual users may wonder how anonymity affects their appearance in the eyes of other users with regard to credibility, trustworthiness, and eventually their messages' persuasiveness.

While we suggest a link between (sender) anonymity and (message) persuasiveness based on the notion of credibility and deindividuation theory (Festinger et al., 1952; Reicher et al., 1995), the literature suggests that social presence may lend itself as an explanatory, mediating factor in this context too (Gefen and Straub, 2004). We incorporate both aspects within a basic mediation model. Our research model is summarized in Figure 1. We develop our research hypotheses in more detail in the following subsections.



**Figure 1: Research Model**

## 2.1 The impact of User Anonymity on Social Presence

Social presence has been defined as a subjective, cognitive measure of how much people develop a feeling that there is a personal, sociable, and sensitive human contact using a certain communication medium (Short et al., 1976). It is often argued that the degree of social presence depends on intimacy and immediacy. The first is related to shared topics of conversations, physical proximity, and eye-contact. Immediacy can be established through verbal and nonverbal cues (Tu and McIsaac, 2002). Yet while it is arguably difficult to evoke human feelings through CMC, it has been shown that social presence can be infused online using socially rich descriptions and pictures (Gefen and Straub, 2004).

This is explained by the social identity model of deindividuation effects (SIDE), which extends self-categorization theory and classical deindividuation theory (Festinger et al., 1952) to focus on situational factors on social influence processes (Reicher et al., 1995). SIDE proposes that under conditions of anonymity, people focus stronger on the perceivable social characteristics, especially those that define a group. Anonymity online (e.g., the absence of names and/or profile photos) reduces the perception of personal and social cues, making it harder for people to develop feelings of social warmth (Postmes and Lea, 2000; Di Blasio and Milani, 2008). Therefore, anonymity may be expected to reduce social presence (Sia et al., 2002). Based on this, our first hypothesis thus states: **H1: Higher levels of anonymity lead to lower levels of social presence.**

## 2.2 The impact of Social Presence on Communication Persuasiveness

Persuasion in CMC has been characterized as an intentional attempt to shape or change behaviors, feelings or thoughts about an issue, object, or action (Fogg, 1998). In CMC, the interaction of argument quality and social comparison processes influence communication persuasiveness (Cotton and Baron, 1980). Social presences directly influences the level of elaborateness in the argument apprehension. However, studies found mixed effects. While some scholars claim that lower social presence leads to focus on the quality of the arguments presented (Short et al., 1976; Sia et al., 2002; Di Blasio and Milani, 2008), others argue that high social presence would enhance argument apprehension (Nunamaker et al., 1991).

Over the course of extended online communication processes, members of a discussion build social norms and become more susceptible to group influence (Postmes et al., 2001; Gefen and Straub, 2004). Social presence facilitates this process as it makes the other group members more salient. Hence, people are more likely to accept the group's opinions and attitudes and may consider its statements more convincing (Postmes et al., 2001). Prior research has also considered social presence in the closely related field of e-learning, where social presence of teachers was found to fostered the students' cognitive processes (Clark and Mayer, 2011), usually associated with increased levels of receptivity. In the domain of e-commerce, Gefen and Straub (2004) linked the existence of social presence on shopping websites to trust in the site operators. The authors argued that social presence builds trust by developing a personal, sociable, human feeling through the website. They found that if a website conveys high social presence, users rather judge the service provider as trustworthy, in particular in terms of integrity. This principle may extend to GDSS communication, where a user benefits from the social presence conveyed by her profile – lifting the trustworthiness, and hence in turn also the persuasiveness of her messages. Based on this, we derive our second hypothesis as follows: **H2: Higher levels of social presence of the discussion participants lead to higher levels of communication persuasiveness.**

### 2.3 The impact of User Anonymity on User Credibility

In line with previous research we define user credibility as the level of trustworthiness and expertise associated with a user (Pornpitakpan, 2004). In their seminal paper, Hovland et al. (1953) investigated persuasive communication by asking: “Who says what to whom with what effect?” Users in anonymous communication settings can answer the “what” part fairly easily as they are able to examine content and quality of a message. Nonetheless, both sender and receiver (i.e. “who” and “whom”) are not transparent. However, source identification can be of great value for the perception of credibility and trust (Fogg et al., 2001). One popular explanation for this is connected to the halo effect, which is a cognitive bias whereby the overall evaluation of another person alters feelings and thoughts about that person’s attributes (Nisbett and Wilson, 1977). People derive personality traits, trustworthiness, and competence from the appearance of another person (Duarte et al., 2012). The halo effect also applies to physical appearance. Positive characteristics such as attractiveness lead to the perception that other positive characteristics are present as well. This has two important implications. First, as users in almost any online platform can choose their profile picture freely and are likely to do so with the intention to appear in a good light, online representations can be assumed to be biased in this regard. The absence of anonymity in online communication may hence be expected to have an overall positive effect on user credibility. Second, given that the provision of a profile picture is possible, its absence may evoke suspicion (Rains, 2007). This is consistent with results from e-commerce research on user reviews, where information from identified sources has been found to be perceived as more useful and more credible (Racherla and Friske, 2012). We hence hypothesize: **H3: Higher levels of anonymity lead to lower levels of user credibility.**

### 2.4 The impact of User Credibility on Communication Persuasiveness

In an extensive meta-study, Pornpitakpan (2004) found that sources of high credibility were consistently considered more persuasive than those with low credibility. This is in line with predictions of leading theories in communication sciences, such as the Elaboration Likelihood Model (ELM, Petty and Cacioppo, 1986). The model proposes differences in persuasion due to how a message is elaborated by an individual. While some messages are processed through a “central route” that enables careful reasoning and evaluation, others take the “peripheral route.” On the latter, factors not linked to the message itself become more salient such as the sender’s attractiveness. For a message to be processed through the central route, individuals need to be mentally motivated and arguments need to be convincing. Then, opinion change is less pronounced as individuals examine arguments more closely (Di Blasio and Milani, 2008). However, if not motivated, cognitive elaboration is reduced and people might be persuaded more easily. Thus, persuasion might happen due to peripheral cues and heuristics. The sender’s general credibility might serve as such a cue. Moreover, this notion supports the idea of halo effects influencing credibility and, essentially, message persuasiveness.

In summary, credible users can be characterized as more persuading than less credible users. We suggest this effect to extend to the messages originating from these users. Our fourth hypothesis hence states that: **H4: Higher levels of user credibility lead to higher levels of communication persuasiveness.**

Taken together, we suggest that the effect of user anonymity on communication persuasiveness is mediated by the factors social presence and user credibility, as depicted in Figure 1.

### **3 Study Design**

In this section, we outline our study design for addressing the research model as presented in Section 2. We intend to use an online survey. Participants are recruited from a pool of registered volunteers at our institution. Participation is incentivized by a gift card lottery among all completed surveys. We present a set of discussion threads. Specifically, we choose topics which (potentially) stimulate a lively and controversial debate among its readers in at least two sufficiently distinct parties. We then pick two arguments from each side and enrich the mere text with an artificial representation of the respective author. After reading and seeing this, the participants state their evaluations with regard to our measures in an online questionnaire.

#### **3.1 Stimuli Material**

After careful rating and discussion among coders, we retrieve articles from a popular German weekly newspaper's website. In order to replicate a group decision setting, we re-frame these texts so that participants would perceive them as business-related discussion task. Second, we identify two opposing views and extracted three comments for each side from the articles' comment section.

#### **3.2 Treatment Design**

We apply a between subjects treatment design. We have four groups that see either the pro or the contra side of a discussion, of which either all comment authors are identified by a photograph or depicted by an anonymous figure.

#### **3.3 Measures**

Whenever possible, and to ensure content validity, previously validated scales are used and adapted to the context of this study. Perceived social presence and credibility items are adapted from Gefen and Straub (2004). All items are measured on a 7-point Likert scale. We measure communication persuasiveness by means of a questionnaire one week prior to the survey that asks for participants' opinion on a range of topics, including those that will be presented in the final survey. This allows us to track possible opinion changes (Stiff, 1993, p. 33). Further constructs tested include hedonic value, perceived discussion quality, and the user's willingness to actively engage him- or herself in the discussion in order to evaluate other factors of interest for GDSS facilitators.

### **4 Discussion and Conclusion**

We expect to observe effects on credibility, social presence and, eventually, persuasiveness due to the variation of user anonymity in our study. First, findings in support of our model would yield further evidence for the devious and immediate effects of anonymity. Such effects would indicate that users wanting to be heard and taken seriously should tend to identify themselves – a process which has previously been observed (Racherla and Friske, 2012). This could prevent discussions from foul language and hoax comments. However, Krasnova et al. (2012) found that in order for user self-disclosure to happen, users need to trust both the provider and the network members and should feel that their privacy is sufficiently protected. Accordingly, considering the short-term effects of anonymity, one could imagine that firms should introduce anonymous communication in GDSS only when necessary, e.g. when discussing delicate issues. The default could be non-anonymity, which allows for an easier assessment of credibility (especially in terms of expertise) and also yields higher social presence. Alternatively, GDSS facilitators should raise user awareness

for possible difficulties in the beginning of a decision making process due to anonymity. It is important to stress that the long-term effects of anonymity have been found to be positive. For instance, groups were found to form social norms when social identities became more salient over time (Postmes et al., 2000; Racherla and Friske, 2012). Hence, negative effects such as the increased number of socially less desirable remarks might disappear. In contrast, full identity in online debates may degenerate the debate towards a dull collection of tame and superficial comments. Conversely, finding no impact of anonymity would nevertheless inform researchers and operators of participatory processes alike.

In summary, this research indicates a path to systematically assess the interplay of user credibility, social presence, and communication persuasiveness for the ever-expanding scope of participatory processes and platforms – both in industrial and municipal contexts.

## 5 References

- Behrendt S, Richter A, Trier M (2014) Mixed methods analysis of enterprise social networks. *Computer Networks* 75(B): 560–577.
- Clark RC, Mayer RE (2011) *E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning*. John Wiley & Sons, New York.
- Cotton JL, Baron RS (1980) Anonymity, persuasive arguments, and choice shifts. *Social Psychology Quarterly* 43(4): 391–404.
- Di Blasio P, Milani, L (2008) Computer-mediated communication and persuasion: Peripheral vs. central route to opinion shift. *Computers in Human Behavior* 24(3): 798–815.
- Duarte J, Siegel L, Young L (2012) Trust and credit: the role of appearance in peer-to-peer lending. *Review of Financial Studies* 25(8): 2455–2484.
- Festinger L, Pepitone A, Newcomb T (1952). Some consequences of de-individuation in a group. *The Journal of Abnormal and Social Psychology* 47(2): 382–389.
- Fogg B, Marshall J, Laraki O, Osipovich A, Varma C, Fang N, Paul J, Rangnekar A, Shon J, Swani L (2001) What makes web sites credible? A report on a large quantitative study. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*.
- Fogg BJ (1998) Persuasive computers: perspectives and research directions. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*.
- Gefen D, Straub DW (2004) Consumer trust in b2c e-commerce and the importance of social presence: Experiments in e-products and e-services. *Omega* 32(6): 407–424.
- Hovland CI., Janis IL, Kelley HH (1953) *Communication and persuasion; psychological studies of opinion change*. Yale University Press, New Haven.
- Krasnova H, Veltri NF, Günther O (2012) Self-disclosure and privacy calculus on social networking sites: The role of culture. *Business & Information Systems Engineering*, 4(3): 127-135.
- Nisbett RE, Wilson TD (1977) The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology* 35(4): 250–256.
- Nunamaker JF, Dennis AR, Valacich JS, Vogel D, George JF (1991) Electronic meeting systems. *Communications of the ACM* 34(7): 40–61.

- Petty RE, Cacioppo JT (1986) The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology* 19: 123–205.
- Pornpitakpan C (2004) The persuasiveness of source credibility: A critical review of five decades' evidence. *Journal of Applied Social Psychology* 34(2): 243–281.
- Postmes T, Lea M (2000) Social processes and group decision making: anonymity in group decision support systems. *Ergonomics* 43(8): 1252–1274.
- Postmes T, Spears R, Lea M (2000) The formation of group norms in computer-mediated communication. *Human Communication Research* 26(3): 341–371.
- Postmes T, Spears R, Sakhel R and Groot D (2001) Social influence in computer-mediated communication: The effects of anonymity on group behavior. *Personality and Social Psychology Bulletin* 27(10): 1243–1254.
- Racherla P, Friske W (2012) Perceived usefulness of online consumer reviews: An exploratory investigation across three services categories. *Electronic Commerce Research and Applications* 11(6): 548–559.
- Rains SA (2007) The impact of anonymity on perceptions of source credibility and influence in computer-mediated group communication: A test of two competing hypotheses. *Communication Research* 34(1): 100–125.
- Reicher SD, Spears R, Postmes T (1995) A social identity model of deindividuation phenomena. *European Review of Social Psychology* 6(1): 161–198.
- Sæbø Ø, Rose J, Molka-Danielsen J (2010) eParticipation: Designing and managing political discussion forums. *Social Science Computer Review* 28(4): 403–426.
- Short J, Williams E, Christie B (1976) *The social psychology of telecommunications*. John Wiley, New York.
- Sia CL, Tan BC, Wei KK (2002) Group polarization and computer-mediated communication: Effects of communication cues, social presence, and anonymity. *Information Systems Research* 13(1): 70–90.
- Tu CH, McIsaac M (2002) The relationship of social presence and interaction in online classes. *The American Journal of Distance Education* 16(3): 131–150.





# A framework for value-based pricing in aircraft surplus markets

Clemens Wickboldt<sup>1</sup> and Natalia Kliewer<sup>1</sup>

<sup>1</sup> Freie Universität Berlin, Professur für Wirtschaftsinformatik, clemens.wickboldt@fu-berlin.de

## Abstract

Aircraft Maintenance, Repair and Overhaul (MRO) companies are facing pricing related decision problems on a daily basis. Although trading platforms are listing offers for some parts, prices are rarely published. Knowing a Fair Market Value (FMV) in this market represents a significant competitive advantage, be it to assess an aircraft for a tear-down or support internal part related decision-making processes. The goal of this work is to develop a framework which is able to aggregate relevant data from different sources, automatically determine a FMV and present context information around it in a transparent and user friendly manner. The Fair Market Evaluator is embedded in a comprehensive software system which enables the user to automatically evaluate single parts or whole part packages. Incomplete historical data sets, aggregation of information from different data sources and integration into existing IT infrastructure are some of the challenges we will address in our ongoing research work.

## 1 Introduction

In order to maintain an airline's fleet, parts are needed in a timely manner. New parts can be procured directly from a supplier, used ones from the aircraft surplus market. If there is no direct use, parts will be stored. Parts with decreasing internal demand will be sold again on the surplus market, which has been growing over the last couple of years. One reason for this growth is an increasing number of parts from aircraft tear-downs. At the same time, this market still lacks transparency, e.g. transactions are usually done via e-mail or telephone. The market for aircraft surplus parts is characterized by a great number of different products but just a few market participants.

There are several ways to determine a price, including auction based market places and prediction models based on historical data. The compartmentalized structure of traders and MRO companies in the aircraft industry may be one reason why there is still no end-to-end auction based market place for surplus parts. As soon as context based FMVs are needed for internal decision making, external determined prices fail to include company-specific information and don't provide a sufficient foundation for decision-making. Predicting a FMV based on historical data however may work well for market players with high amounts of structured and high quality historical data. The

goal of this work is to develop a framework which is able to automatically determine a FMV and present context information around it in a transparent and user friendly manner.

An example use case would be: A supplier offers an airplane for a tear-down. The value of this plane would be the sum of all the parts in it. The user would have to determine a Fair Market Value for every single part of this plane and sum it all up to get an idea of what would be a reasonable price for it. Alternatively, the user could let the Fair Market Evaluator calculate values for each item and in that way provide a solid foundation for decision making. If there is not enough data available for a reliable Fair Market Value calculation, the user still gets all the aggregated information as internal historical prices and external market prices, which supports a manual Fair Market Value determination.

## 2 Fair Market Value-based pricing

In order to find an approach to determinate a Fair Market Value, we take a look at existing pricing methods, define the term *Fair Market Value* and see if we can combine both.

It has been shown that price has the greatest impact on a company's earnings before interest and taxes (Hinterhuber 2004). The way a product's or service's price is determined can be categorized into three different groups (Hinterhuber 2008):

1. Cost-based pricing uses data from cost accounting and for example considers original purchase price and internal costs (e.g. repairing) to determine a sale price. It doesn't take the competition or market into account and is therefore considered as the weakest approach.
2. Competition based-pricing observes price levels of the competition and uses market prices for orientation in price setting.
3. Value-based pricing uses a predefined value as the basis for determining the price.

As we are looking for a value which is able to serve as foundation for decision making and price setting, value-based pricing seems the way to go. In the literature, value is defined as the difference between the customer's anticipated benefit of a good or service and the supplier's cost of producing it (Johansson et al. 2015) or as benefits received in exchange for price (Anderson & Narus 1998). This work focuses on a business to business secondary market in the aircraft industry, which mainly differentiates between the following value terms:

- Base Value is considered as the economic value and assumes balanced supply and demand (Kelly 2008), completely informed market participants and is considered as a hypothetical value (Ackert 2012).
- (Current) Market Value is defined as determined value after manual analysis (Ackert 2012) or the "most likely trading price" and is synonymous with Fair Market Value (Kelly 2008).

Still, many companies are not able to benefit from value-based pricing. One obstacle is that the value has to be determined before it can be used as an argument for pricing (Hinterhuber 2008). Value-based pricing and value estimation methods are often related to product introductions in primary markets. State of the art literature about value-based pricing takes the customer point of view for value determination. Common methods are surveys or conjoint analysis (Anderson et al. 1992). This is applicable to this scenario in a limited way only. Our sample data contains more than

500.000 different parts<sup>1</sup>. Conducting surveys for this large quantity of different products seems to be a disproportionate effort. Also, as soon as context-based values are needed, external determined prices fail to include company-specific information and don't provide a sufficient foundation for a company's decision-making. Therefore, a new and innovative approach is needed.

We are working on a framework which is able to determine a part's value under consideration of the company's specific situation. It contains information about the following:

- Original launch price,
- historical purchase price within the company,
- historical offer price within the company,
- historical selling price within the company,
- historical internal and manual evaluation and
- current external market price.

We suggest a combination of a set of business intelligence methods to transparently determine and present a Fair Market Value. Based on the results of Voß & Lessmann (2013), who had determined prices for the used car market, an application of prediction methods seems to be promising to evaluate secondary market items. There is a wide variety of methods<sup>2</sup>, which may perform differently in this specific scenario and will be evaluated in the final paper.

The Fair Market Evaluator framework delimits from state of the art value-based pricing methods as shown in table 1. Value-based pricing and value estimation methods are related to product introductions in primary consumer markets.<sup>3</sup> This work focuses on a business to business secondary market in the aircraft industry. Value-based pricing focuses on an individual consideration of a small amount of products, which is an obvious scope for pricing related customer surveys and conjoint analysis. This work aims at automatic evaluation of thousands of different products and still present the valuation in a transparent way.

Dimension	State of the art	Fair Market Evaluator
Target	Primary Markets	Secondary Market
Audience	B2C	B2B
Scope	Small amount of products	Thousands of products
Automation	Individual consideration	Automated evaluation

**Table 1: Delimitation from state of the art value-based pricing methods**

An automated estimation of the FMV as a basis for evaluating purchasing and selling prices represents a promising approach to address the transparency issue. It could not only be used for supporting the surplus part utilization process but also in any department which would work with value-based pricing methods.

<sup>1</sup> Real world data is being provided by a large MRO company and can be used to test the Fair Market Evaluator.

<sup>2</sup> See Hastie et al. (2009) for an overview.

<sup>3</sup> See for example Hinterhuber (2004, p. 767).

### 3 Fair Market Evaluator embedded in comprehensive software system

After defining the approach to determine the Fair Market Value, we introduce a software framework which aggregates information from different sources, prepares data for the Fair Market Evaluator model and presents the results to the end user.

In the past, FMV determination has been done manually, mainly based on experts' estimates, including the risk of not using all available information to determine the FMV objectively. Until now and for various organizational reasons, not all historical pricing data is collected and stored. Because of incomplete datasets the FMV can be recognized directly in only a few cases. Prediction models naturally assume a complete set of historical data and can't deal with missing datasets. We will address this challenge by relying on statistical methods for dealing with missing data.

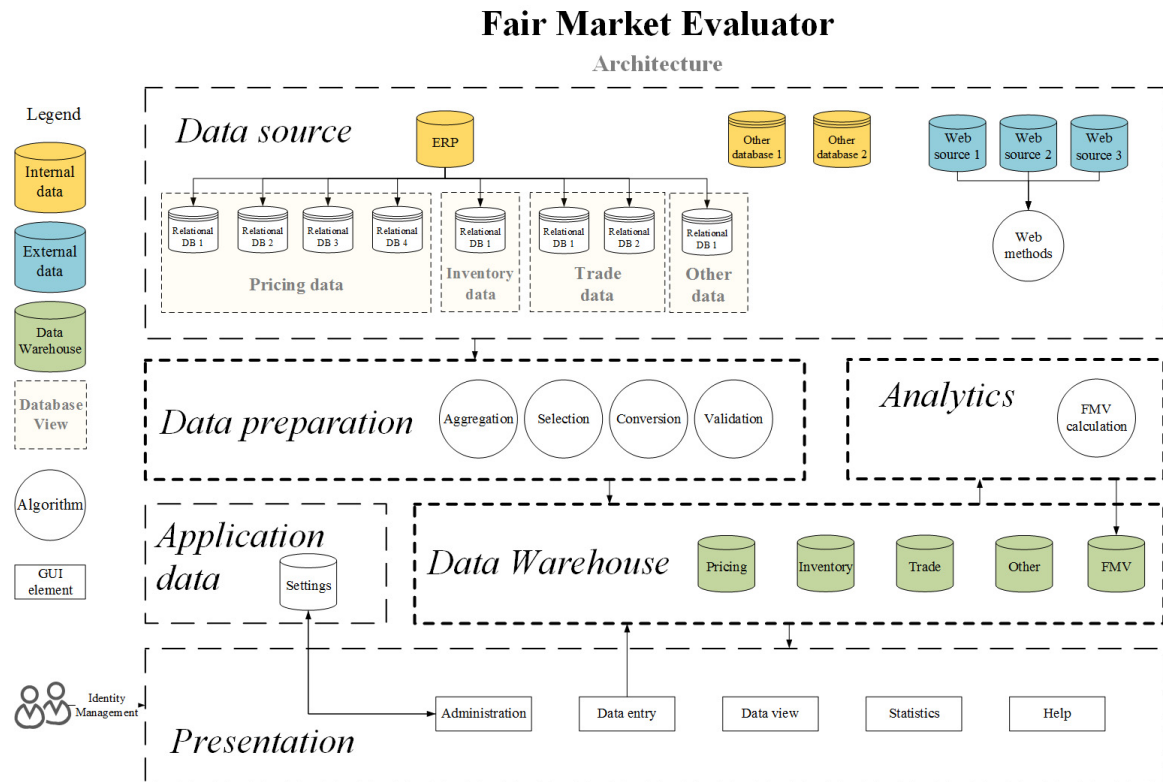
A similar approach is a web marketing system with automatic pricing, called *digiprice*, which was proposed by Abe and Kamba (2000). Their goal is to maximize the profit within an on-line marketing site. Table 2 shows the differences between our proposed Fair Market Evaluator and *digiprice*.

Dimension	digiprice	Fair Market Evaluator
Market	Digital content (p. 775)	Aircraft surplus parts
Audience	'consumers' (p. 776)	B2B
Scope	Marketing and pricing (p. 775)	Value determination
Goal	Find price which maximizes the overall profit (p. 775)	Find value which reflects the part's value within the company
Considered variables	Historic prices, historic sales, product description provided by seller (p. 776)	Introductory price, current market price, historic purchase price, repair costs, historic offer, historic sell price, historic manual evaluation
Algorithms	Stochastic and linear approximation of optimal price (p. 779 and p. 781)	Forecasting per multiple linear regression for now.

**Table 2 - Differentiation of Fair Market Evaluator from alternative framework**

As *digiprice* is looking to price digital content for consumers, we're determining a value for aircraft surplus parts in a B2B market. Also, we're not looking for a sales price only but for a value which is the foundation for decision making and a basis for deducing prices. At last, we consider multiple dimensions within a linear regression model whereas *digiprice* considers three dimensions and uses stochastic and linear approximation methods.

The Fair Market Evaluator itself is embedded in a comprehensive software system to support pricing, trading and logistic processes, see figure 1 for an approach.



**Figure 1: Approach for a Fair Market Evaluator system architecture**

The system architecture of the Fair Market Evaluator is divided into six layers. The first layer contains data gathering from different sources. Internal data lies in a corporate ERP with relational databases in the back-end, which are aggregated in logical database views for convenient access. Additionally, there are other (non-)relational databases, outside the ERP. External data is collected via web interfaces from data providers, e.g. trading platforms. In the data preparation layer, data is aggregated, relevant data is selected, converted to one consistent format and validated in the end. The data flows in one direction to avoid contamination of the production data. The Data Warehouse layer contains one data base per data category. In this case, there are pricing, inventory, trade and master data. The analytics layer calculations are supplied by Data Warehouse data. Calculated FMVs are stored in a database in the Data Warehouse layer. The presentation layer provides the user interface. The user is able to enter data into the data warehouse as well as change existing data. In this way, the user is able to override a model-suggested Fair Market Value which will enrich the model's quality with information about the prediction's accuracy. The presentation layer also provides statistics which are only available because of the prepared data stored in the Data Warehouse. Administrative settings are stored in a database within the application data layer. Authentication is handled by a corporate identity management system.

## 4 Outlook

An automated FMV calculation may deliver valid results if the underlying data is complete and of good quality. An automated value determination addresses the market characteristic of many different products. Data aggregation prior to the Fair Market Value calculation could be one way to build a generic surplus part trading platform and overcome the market's transparency issue. In the

end, an automated value determination prevents bad part utilization decisions and could lead to a more efficient and sustainable use of surplus material.

Subsequent to introducing the Fair Market Evaluator framework, an evaluation of its utility and benefit will follow. A possible approach to determine the prediction assurance ex-post is to measure the amounts of FMVs which had been overwritten by the user. Evaluation of the framework in practice and empirical insights in cooperation with a MRO company are one part of the final paper.

## 5 References

- Abe N, Kamba T (2000). A Web marketing system with automatic pricing. *Computer Networks- the International Journal of Computer and Telecommunications Networking*, 33(1-6), 775-788.
- Ackert S (2012). Basics of aircraft market analysis. *Aircraft Monitor*
- Anderson JC, Jain DC, Chintagunta PK (1992). Customer value assessment in business markets: A state-of-practice study. *Journal of Business-to-Business Marketing*, 1, 3-29.
- Anderson JC, Narus JA (1998) Business marketing: understand what customers value. *Harvard business review*, 76, 53-67
- Hastie T, Tibshirani R, Friedman J (2009) *The elements of statistical learning: Data mining, inference, and prediction*. Springer Series in Statistics (2)
- Hinterhuber A (2004) Towards value-based pricing - An integrative framework for decision making. *Industrial Marketing Management*, 33, 765-778
- Hinterhuber A (2008) Customer value-based pricing strategies: why companies resist. *Journal of business strategy*, 29 (4), 41-50
- Hinterhuber A, Bertini M (2011) Profiting when customers choose value over price. *Business strategy review*, 22 (1), 46-49
- Johansson M, Keränen J, Hinterhuber A, Liozu S, Andersson L (2015) Value assessment and pricing capabilities - how to profit from value. *Journal of Revenue & Pricing Management*.
- Kelly DB (2008). Forecasting aircraft values: An appraiser's perspective. *Air-finance annual*.
- Voß S, Lessmann S (2013). Resale price prediction in the used car market. *Tristan Symposium VIII*

# Vorstudie zur Digitalisierung analoger Medienprodukte – Eine quantitative Datenanalyse anhand der Branche Brettspiele

André Witzel<sup>1</sup> und Jürgen Karla<sup>1</sup>

<sup>1</sup> Hochschule Niederrhein University of Applied Sciences, andre.witzel@stud.hn.de, juergen.karla@hs-niederrhein.de

## Abstract

Digitalisierung ist eine der großen Herausforderungen der Wirtschaftsinformatik. Dennoch steht oftmals nicht die Frage im Raum, ob etwas digitalisiert werden kann, sondern ob sich die Umsetzung überhaupt lohnt. Dies resultiert in der Fragestellung, wie ein Potential zur Digitalisierung analoger Produkte festgestellt werden kann. Im Rahmen des vorliegenden Beitrags wird diese Fragestellung bezogen auf das analoge Produkt Brettspiel und dessen digitaler Umsetzung in Apps für mobile Endgeräte untersucht. Dazu wird im Rahmen einer ersten Analyse dieser ausgewählte Bereich der analogen Medien näher betrachtet und statistisch analysiert. Als Fundament der Untersuchung dienen Daten, die aus frei zugänglichen Quellen erhoben, aufbereitet und verknüpft wurden. Die Ergebnisse zeigen, dass erste Aussagen für dieses Marktsegment getroffen werden können und damit eine weitere Untersuchung der Daten für Wirtschaft und Wissenschaft vielversprechend ist.

## 1 Einleitung

Digitalisierung ist in unser alltägliches Leben eingezogen und verändert es grundlegend. Sie ist ein bedeutender Trend für Wirtschaft, Gesellschaft sowie Politik und damit eine der großen Herausforderungen der Wirtschaftsinformatik (Hanelt et al. 2015; Piccinini et al. 2015; BMBF o.J.). Betroffen sind insbesondere Hersteller und Anbieter analoger Medien. Sie müssen zeitnah dem Trend folgen, um am Markt nicht gegen die Konkurrenz zu verlieren; dürfen aber gleichzeitig nicht überstürzt handeln, um Fehler zu vermeiden, die die eigene Existenz bedrohen könnten. Dabei stellt sich das Umsetzungspotential ausgewählter analoger Medien in digitale Produkte durchaus unterschiedlich dar. Bekannt ist dieses Phänomen u.a. aus der Printmedienbranche (Schäfer 2002). Doch neben Ansätzen wie *blendle* (blendle o.J.; Müller 2015), die das *Wie* der Digitalisierung lösen, stellt sich die Frage *Was* digitalisiert werden sollte. D.h., welche der analogen Produkte mehr oder weniger Potential für eine digitale Umsetzung bieten und vor allem, wie ein solches Potential ermittelt werden kann.

Ein besonderer Treiber der Digitalisierung sind mobile Endgeräte – Smartphones und Tablets. Mit der Einführung des iPhone im Jahr 2007 begann ein gravierender Wandel hin zum digitalen Freund in der Tasche und mit ihm der Boom der mobilen Applikationen (Apps) (Buck et al. 2014). Insbesondere Mobile Games sind einer der wesentlichen Treiber des Mobile Business (Basole, Karla 2012). Ein analoges Pendant hierzu sind klassische Brettspiele. Der Brettspielmarkt ist in Deutschland ein etablierter und stabiler Markt (n-tv 2014) und zeigt in den USA und in Kanada ein großes Wachstumspotential auf (ICv2 2015). Es liegt daher nahe, die Ausprägungen der Digitalisierung in diesem Marktsegment zu untersuchen.

Im vorliegenden Beitrag soll anhand der ausgewählten Branche der Brettspiele eine deskriptive Analyse der Umsetzungen in digitale Angebote durchgeführt werden. Dabei liegt der Fokus auf einer Merkmalsanalyse, welche zur Erhebung eines Umsetzungspotentials herangezogen werden soll.

## 2 Vorgehensweise

Die Vorgehensweise teilt sich grob in drei Schritte. Zunächst wurden die für die Untersuchung notwendigen Daten aus frei zugänglichen Datenquellen erhoben. Im Anschluss wurden die Daten zwecks Verknüpfung aufbereitet. Darauf aufbauend wurde eine Analyse mittels ausgewählter Verfahren der deskriptiven Statistik durchgeführt.

### 2.1 Datenerhebung

Primär wurden die der Analyse zugrunde liegenden Daten aus zwei Quellen gewonnen: *BoardGameGeek* als Quelle für die Brettspiel-Daten und *App Annie* als Quelle für die App-Daten. Sämtliche Daten wurden zwischen Juli 2015 und August 2015 erhoben.

#### 2.1.1 Erhebung der Brettspiel-Daten: BoardGameGeek

*BoardGameGeek (BGG)* (BGG o.J. a) wurde im Jahr 2000 als freie Brettspiel-Online-Datenbank und Community gegründet. Die Datenbank enthält über 80.000 gelistete Spiele (Stand Oktober 2015) und ist damit die bedeutsamste Brettspiel-Datenbank der Welt. Neben den reinen Datenbank-Funktionen bietet die Website ihren Nutzern u.a. die Möglichkeit eigene Spielereihen zu erstellen, Spiele zu bewerten oder sich in Foren auszutauschen. *BGG* erfreut sich großer Beliebtheit in der Community und ist der zentrale internationale Anlaufpunkt für Hobbyisten und Sammler (BGG o.J. b; BGG o.J. c; BGG o.J. d).

Zunächst wurden aus der Grundgesamtheit der möglichen Spiele 469 Spiele ausgewählt. Die Stichprobe kam zu ungefähr gleichen Teilen durch eine bewusste sowie eine willkürliche Auswahl zustande. Bei der bewussten Auswahl wurden Spiele ausgewählt, von denen bekannt ist, dass eine Umsetzung als App bereits existiert(e).

Zu den ausgewählten Brettspielen wurden für die spätere Analyse u.a. folgende Daten erhoben:

- *Primary Name*: Der meistgebrauchte Name bzw. der originäre Handelsname;
- *Alternate Name*: Eine Liste alternativer Namen;
- *Year Published*: Das Veröffentlichungsdatum;
- *Min. # Players*: Die minimale Anzahl an benötigten Spielern;



- *Max. # Players*: Die maximal erlaubte Anzahl an Spielern;
- *Playingtime*: Die durchschnittliche Spielzeit, die benötigt wird, um das Spiel genau einmal nach den vorgeschriebenen Regeln zu beenden;
- *Min. Age*: Das vorgeschlagene minimale Alter, um das Spiel zu spielen;
- *Avg. Rating*: Der Durchschnitt sämtlicher Bewertungen der Nutzer von BGG auf einer Skala von 1 (schlecht) bis 10 (gut);
- *Avg. Game Weight*: Der Durchschnitt sämtlicher Bewertungen des Anspruchsniveaus des Brettspiels der Nutzer von BGG auf einer Skala von 1 (leicht) bis 5 (schwer);
- *Mechanics*: Eine Liste von Spielmechaniken, die in dem Spiel vorkommen.

Die Datenerhebung erfolgte mittels eines Java-Skripts, unter Rückgriff auf die durch *BGG* angebotene API (BGG o.J. e; BGG o.J. f). Die erhobenen Daten wurden zur weiteren Aufbereitung in eine Tabellenkalkulationssoftware abgelegt.

### 2.1.2 Erhebung der App-Daten: App Annie

*App Annie* ist ein 2010 gegründetes Analyse- und Business-Intelligence-Unternehmen. Das Unternehmen bietet auf seiner gleichnamigen Plattform (*App Annie* o.J.) verschiedene Analyse- und Statistik-Tools rund um den Markt mobiler Applikationen (*LinkedIn* o.J.).

Aufgrund der marktführenden Position wurden bei der Erhebung ausschließlich Apples App Store und Googles Play Store in den Untersuchungsbereich einbezogen (*IDC* o.J.). Während im Rahmen dieser Erhebung zwar zunächst die Frage nach einer digitalen Umsetzung der ausgewählten Brettspiele im Vordergrund stand, so wurden für weiterführende Untersuchungen bereits Regeln zur Auswahl der Apps angewandt:

1. Gibt es eine inoffizielle und offizielle Version, so wurde die offizielle gewählt;
2. gibt es eine freie und eine bezahlte Version – d.h. im Regelfall eine Demo und eine Vollversion – so wurde die Bezahlversion gewählt;
3. keine Unterscheidung zwischen Smartphone- und Tablet-Apps;
4. Umsetzungen, die zwischenzeitlich nicht mehr verfügbar sind, gelten als umgesetzt.

Folgende Daten wurden u.a. je App Store erhoben:

- *Name*: Name der Anwendung;
- *Price*: Preis zum Erwerb der Applikation in Dollar;
- *In-App-Purchases*: Binäres Merkmal, ob In-App-Käufe existieren;
- *Average Rating*: Bewertung der Applikation auf einer Skala von 1 (schlecht) bis 5 (gut).

Mögliche Umsetzungen wurden auf der Plattform von *App Annie* anhand des *Primary Name* und ggfs. anhand der *Alternate Names* manuell gesucht.

## 2.2 Aufbereitung und Analyse der Daten

Nach erfolgreicher Erhebung wurden die Daten für die Analyse aufbereitet und mittels deskriptiver Verfahren analysiert.

Bereits bei der Erhebung wurden die *Mechanics* zu einem dichotomen Merkmal mit den Ausprägungen 0 – für nicht vorhanden – und 1 – für vorhanden – transformiert. Im selben Zuge wurde die Anzahl der Mechaniken (*# Mechanics*) bestimmt.

Anschließend wurden die Brettspiel-Daten und App-Daten gefiltert und zu einer einheitlichen Tabelle zusammengefasst. Dabei wurden fast alle zuvor genannten Daten aus BGG unverändert übernommen. Die Ausnahmen sind:

- *Alternate Name*: Die Liste der Namen war nur für die Erhebung relevant;
- *Avg. Rating*: Das Rating wurde durch 2 geteilt und auf eine Nachkommastelle gerundet, um der Skala des Average Rating der App Stores gegenüber gestellt werden zu können;
- *Avg. Game Weight*: Wurde auf eine Nachkommastelle gerundet.

Weiterhin wurden die Daten aus App Annie zu neuen Merkmalen transformiert, wodurch sich u.a. folgenden Merkmale ergeben:

- *App*: Trinärer Wert, ob eine App-Umsetzung zu dem Spiel existiert (0 – keine; 1 – Apple oder Google; 2 – Apple und Google);
- *Shop Rating*: Mittelwert der *Average Ratings* der App Stores, falls vorhanden.

Eine Betrachtung des *Price* und der *In-App-Purchases* wurde zunächst nicht weiter verfolgt, da diese bestimmten Geschäftsmodellen obliegen, die zunächst nicht im Fokus der Betrachtung liegen (s. Kap. 4).

Im nächsten Schritt wurden für jedes Merkmal die Häufigkeiten, Minimum, die Quartile inkl. Median, Maximum, der arithmetische Mittelwert, die Standardabweichung sowie der Variationskoeffizient ermittelt. Dies geschah sowohl für die gesamte Menge (*Alle*) der Stichprobe, als auch jeweils für die Teilmenge der umgesetzten Brettspiele (*App*) und dessen Komplement von nicht umgesetzten Brettspielen (*keine App*). Eine Ausnahme bildet hierbei das *Shop Rating*, da dies ausschließlich für umgesetzte Brettspiele Ausprägungen hat.

Da der Mittelwert der *# Mechanics* bei der Menge *App* bei 2,76 liegt – die digital umgesetzten Brettspiele also im Mittel zwischen 2 bis 3 Mechaniken nutzen – wurden die auftretenden Kombinationen zwei- und dreidimensional in Kreuztabellen ausgewertet.

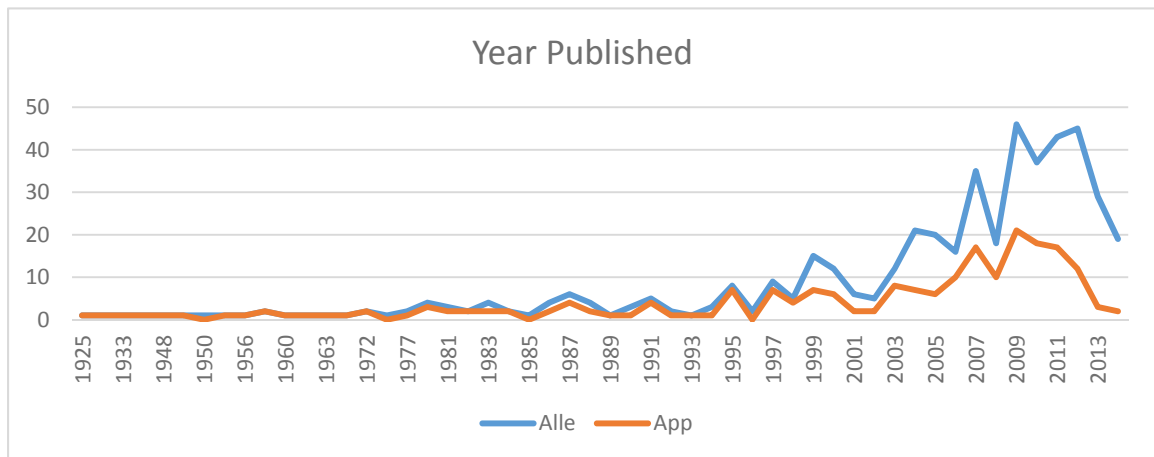
Zur Visualisierung wurden verschiedene Graphiken erzeugt, u.a. Stabdiagramme für die einzelnen Häufigkeiten sowie Liniendiagramme für den Vergleich der Häufigkeiten.

### 3 Interpretation und Hypothesen

Für ein besseres Verständnis der Interpretation soll zunächst angemerkt werden, dass die Gesamtzahl der betrachteten Brettspiele (*Alle*) die Obergrenze der möglichen Umsetzungen darstellt. Dies bedeutet, dass bei den Ausprägungen, bei denen keine oder nur wenige Brettspiele existieren, keine oder nur eine geringfügige Aussage über ein Potential zur Digitalisierung möglich ist. Auch kann der Anteil der nicht umgesetzten Spiele sowohl (a) eine Aussage darüber sein, ob das Potential verschenkt wurde oder (b) ob die Spiele aus einem bestimmten – nicht zu erhebenden – Grund nicht bzw. schlechter umsetzbar sind. Daher erscheint die Auswahl der Grafik beider Häufigkeiten – *Alle* und *App* – sinnvoll, um gleichzeitig die Verteilung, die Obergrenze und den Abstand darzustellen.

### 3.1 Year Published

Wie in Bild 1 zu erkennen ist, verteilen sich die Brettspiele (*Alle*) vermehrt auf den Veröffentlichungszeitraum nach 1995. Dies reflektiert die Entwicklung des Marktes, welche ab Anfang/Mitte der 1990er Jahre ein starkes Wachstum zeigte – angetrieben u.a. durch die Veröffentlichung von *Die Siedler von Catan* 1994 und später eine zunehmende Internationalisierung. Dabei liegen ca. 50% aller Veröffentlichungen zwischen 2001 und 2011. Die digitalen Umsetzungen (*App*) folgen dem Verlauf mit einer zeitlichen Verzögerung.



**Bild 1: Year Published**

Deutlich sichtbar wird, dass die digitalen Umsetzungen mit einem zeitlichen Verzug realisiert werden und z.B. nicht parallel entwickelt werden. Zwei mögliche Begründungen dafür sind, dass sich der Markt der Apps erst seit ca. 2010 entwickelt (Buck et al. 2014), sowie, dass Spieleverlage warten bis das jeweilige Brettspiel einen gewissen Erfolg erreicht hat.

### 3.2 Min. & Max. # Players

Die Mehrheit der minimalen Anzahl liegt bei zwei Spielern (Bild 2). Bei der maximalen Anzahl ist eine Spitze bei zwei Spielern; die Mehrheit liegt zwischen vier und fünf Spielern (Bild 3). Die digitalen Umsetzungen folgen dem jeweiligen Verlauf, mit der Ausnahme, dass diese bei der maximalen Anzahl der Spieler schneller abflacht.

Bei digitalen Umsetzungen mit mehr als einem Spieler muss entweder eine Möglichkeit der Vernetzung oder des gleichzeitigen resp. abwechselnden Spielens gegeben sein. Alternativ muss eine Künstliche Intelligenz (KI) den Part der Gegenspieler übernehmen.

Der Durchschnitt der minimalen Anzahl an Spielern und die klare Mehrheit liegen bei zwei. Dies zeigt, dass eine digitale Umsetzung mit mehr als einem Spieler keine größere Hürde darstellt. Insbesondere das mobile Internet ermöglicht eine einfache und schnelle Vernetzung der Spieler.

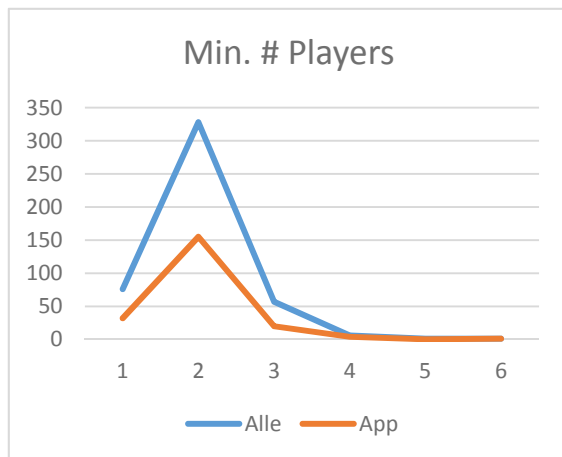


Bild 2: Min. # Players

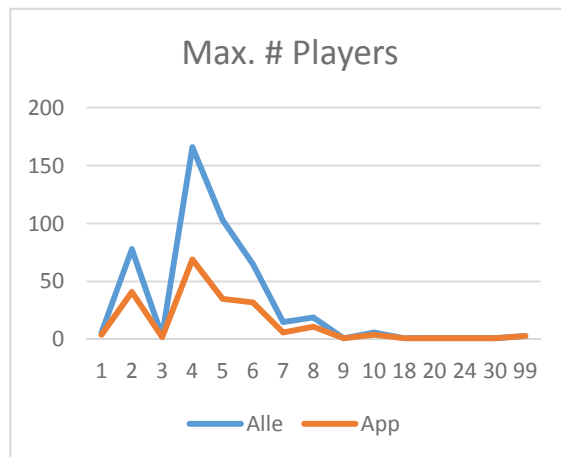


Bild 3: Max. # Players

Die maximale Anzahl von Spielern – mit einer Mehrheit von zwei bis vier Spielern – kann ein Indikator für die akzeptablen Obergrenzen der digitalen Umsetzungen sein. So kann es sein, dass vier Spieler auf einem (entsprechend großen) Gerät noch akzeptabel sind, während sechs Spieler die Obergrenze für die akzeptable Vernetzung darstellen könnte. Je mehr Spieler an einem Spiel teilnehmen können, desto mehr Strategien können eventuell in dem Spiel zielführend sein. Damit muss auch die KI intelligenter werden und diese Strategien kennen und erkennen können. Mehrere Strategien bedeuten auch, dass die digitale Umsetzung selber komplexer werden könnte. Allerdings sollte beachtet werden, dass die maximale Anzahl der Spieler bei analogen Brettspielen oft aufgrund des beiliegenden Materials – einem wesentlichen Kostentreiber der analogen Umsetzung – eingehalten werden muss.

### 3.3 Playingtime

Die durchschnittliche Spielzeit der Brettspiele verteilt sich auf einzelne Spitzen bei 30, 45, 60, 90 und 120 Minuten, was den üblichen Angaben der Spieleverlage auf den Schachteln entspricht. Bei den digitalen Umsetzungen zeichnet sich der Trend zur Umsetzung von kürzeren Spielen ab (Bild 4).

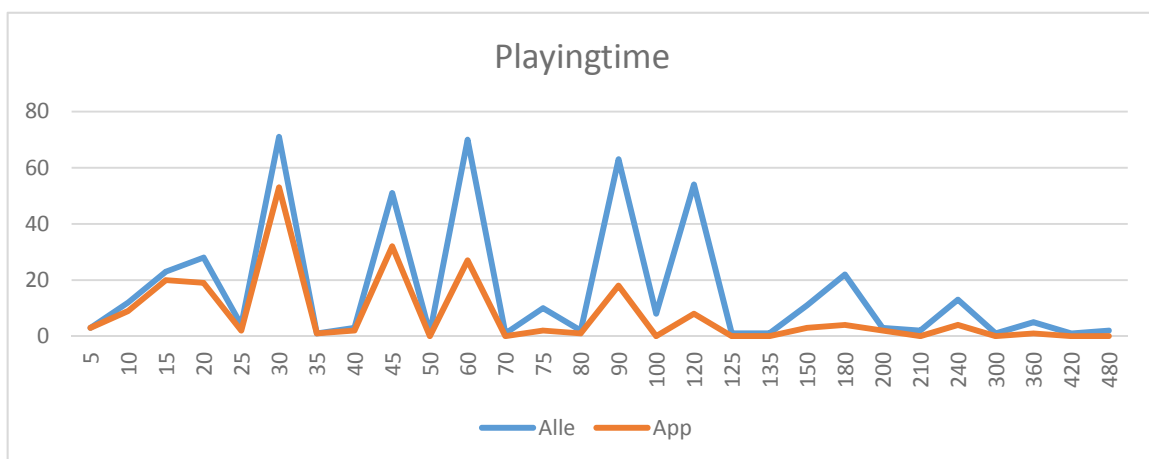


Bild 4: Playingtime

Insbesondere Umsetzungen von Brettspielen mit einer Spieldauer von kleiner gleich 60 Minuten sprechen für Spiele, die in Pausen, Bahnfahrten o.Ä. gespielt werden können bzw. *Casual Games*. Zu kurze Spiele – unter 15-20 Minuten – scheinen genauso unbeliebt zu sein, wie zu lange Spiele – oberhalb der 60 Minuten. In der digitalen Umsetzung muss bei längeren Spielen zwecks möglicher Unterbrechung entweder (a) die Möglichkeit der Speicherung etabliert werden oder (b) ein asynchrones Spiel gegen Online-Mitspieler ermöglicht werden. Zu beachten ist, dass die Spielzeit Bezug auf die analogen Brettspiele nehmen, bei denen durch das Warten auf Mitspieler, Setzen von Steinen, Berechnen von Punkten, etc. ebenfalls Zeit vergeht. Solche Elemente sind bei der digitalen Umsetzung i.d.R. automatisiert und verkürzen entsprechend die Spielzeit.

### 3.4 Avg. Game Weight

Bild 5 zeigt, dass die durchschnittliche Gewichtung sich vornehmlich zwischen 1,1 und 3,7 verteilt. Das untere Quartil liegt bei 1,8, der Median bei 2,4 und das obere Quartil bei 3,0. Im Mittel liegen die Brettspiele ca. bei einer Gewichtung von 2,46. Die digitalen Umsetzungen starten ebenfalls bei 1,1. Erkennbar ist ein leichter Abwärtstrend ab einer Gewichtung von 2,0 und einem rapiden Abfall ab 2,5. Gut 57% aller Spiele haben eine Gewichtung von 2,0 oder weniger. Oberhalb der 3,0 gibt es nur wenige digitale Umsetzungen.

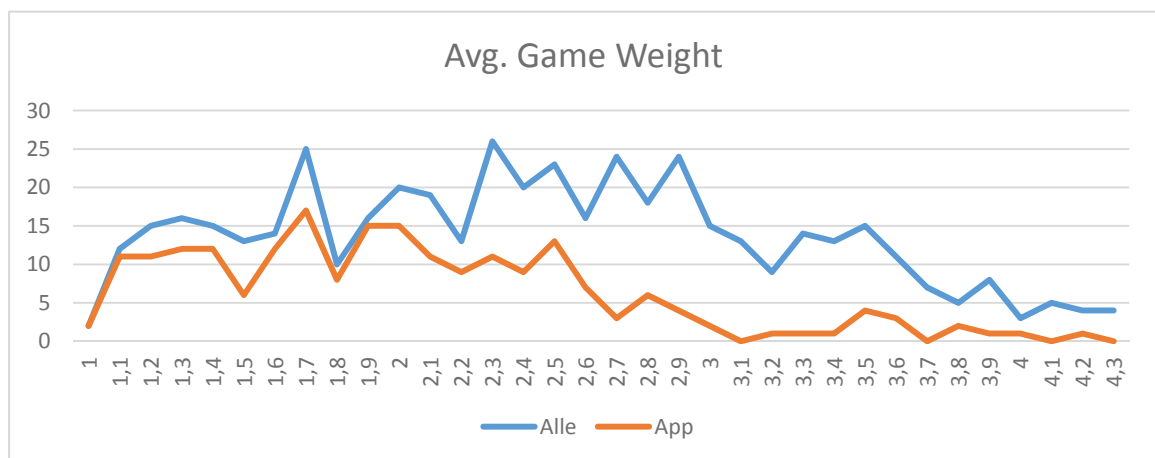


Bild 5: Avg. Game Weight

Im Folgenden sollen zunächst einige Vergleichswerte der durchschnittlichen Spielgewichtung gezeigt werden.

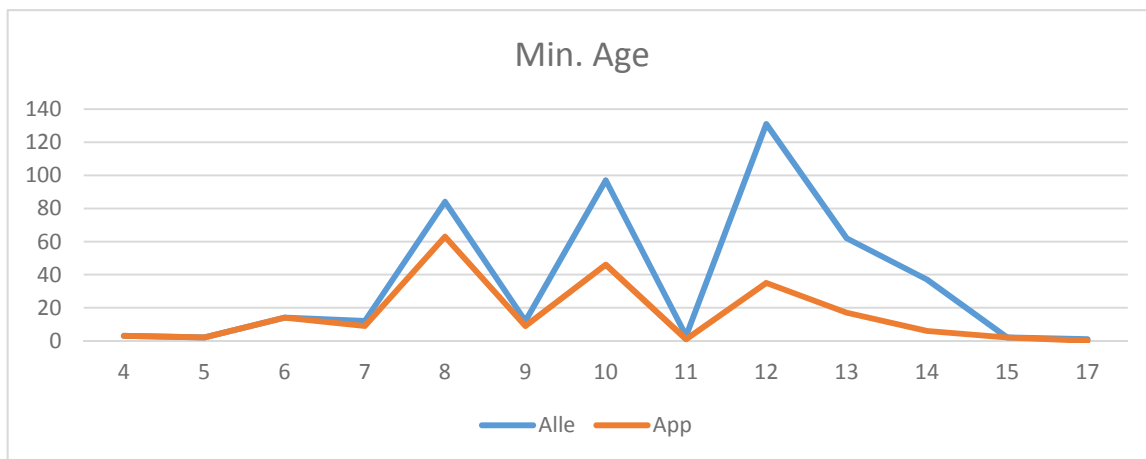
- The Game of Life: 1,2
- The aMAZing Labyrinth: 1,4
- Monopoly: 1,7
- Carcassonne: 2,0
- Scrabble: 2,2
- Mü & More: 2,5
- Diplomacy: 3,4
- Schach: 3,7

Dadurch lassen sich die Brettspiele grob kategorisieren: Bei Brettspielen mit einer Bewertung ab ca. 1,3 bis 2,0 ist ein grundlegendes strategisches Denken von Nöten. Ein bis zwei Züge sollten im Voraus bedacht werden. Brettspiele mit einer Bewertung von 2,0 bis hin zu ca. 2,6/2,7 erfordern bereits ein vorausschauendes Handeln und die Beachtung verschiedener Zusammenhänge. Es scheint, dass seine Bewertung von 3,0 und höher bereits ein Kriterium zur Erkennung strategischer Brettspiele ist. Die durchschnittliche Gewichtung macht damit eine konkrete Aussage über Komplexität und/oder Kompliziertheit des Brettspiels.

Weniger komplexe/komplizierte Spiele – unterhalb der 2,0 – scheinen eher Gegenstand einer digitalen Umsetzung zu sein. Dies und dass die untere Grenze bei 1,1 liegt, sprechen für leichte Spiele in Richtung *Casual Games* (Manzius 2012).

### 3.5 Min. Age

Beim Mindestalter sind klare Spitzen bei 8, 10 und 12 Jahren zu erkennen. Die digitalen Umsetzungen weisen eine Tendenz zu einem geringeren Alter, nicht aber unter 8 Jahre auf (Bild 6).



**Bild 6: Min. Age**

Ob das Mindestalter von Brettspielen – welches u.a. auch aufgrund rechtlicher Vorgaben gesetzt wird – 1:1 auf die digitale Version übertragbar ist, ist zunächst fraglich. Allerdings lässt das Mindestalter auch auf die Schwierigkeit der Inhalte und die Komplexität/Kompliziertheit des Spieles zum Erreichen des Spielziels schließen.

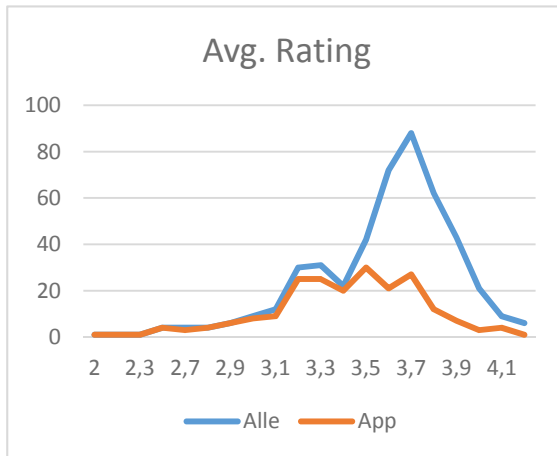
Anhand der Spitzen ist zu erkennen, dass das "optimale Mindestalter" zwischen 8 und 12 Jahren liegt. Dies liegt offensichtlich zum größten Teil an der Selbstkontrolle der Medienbranche (Spielbar o.J.). Allerdings entspricht es auch dem ungefähren Zeitraum, in dem Kinder ihr erstes Smartphone/Tablet oder Zugriff auf ein Gerät der Eltern erhalten (BITKOM 2014).

Die Tendenz zu einem geringeren Alter bei den digitalen Umsetzungen zeigt, dass weniger komplexe/komplizierte Spiele – u.a. *Casual Games* (Manzius 2012) – bevorzugt werden, aber auch, dass Kinder eine beliebte Zielgruppe für den Markt digitaler Spiele sind (Grubb 2015).

### 3.6 Avg. Rating & Shop Rating

Die durchschnittliche Bewertung (Bild 7) verteilt sich hauptsächlich zwischen 3,2 und 3,9, wobei 47,44% aller Brettspiele eine Bewertung zwischen 3,6 und 3,8 haben. Bei den digitalen Umsetzungen haben über 75% aller Spiele relativ gleichverteilt eine Bewertung zwischen 3,2 und 3,8.

Bild 8 zeigt die Verteilung des Avg. Rating und des Shop Rating übereinandergelegt. Zu erkennen ist, dass das Shop Rating vergleichsweise besser ausfällt mit einer Mehrheit zwischen 3,6 und 4,4.



**Bild 7: Avg. Rating**

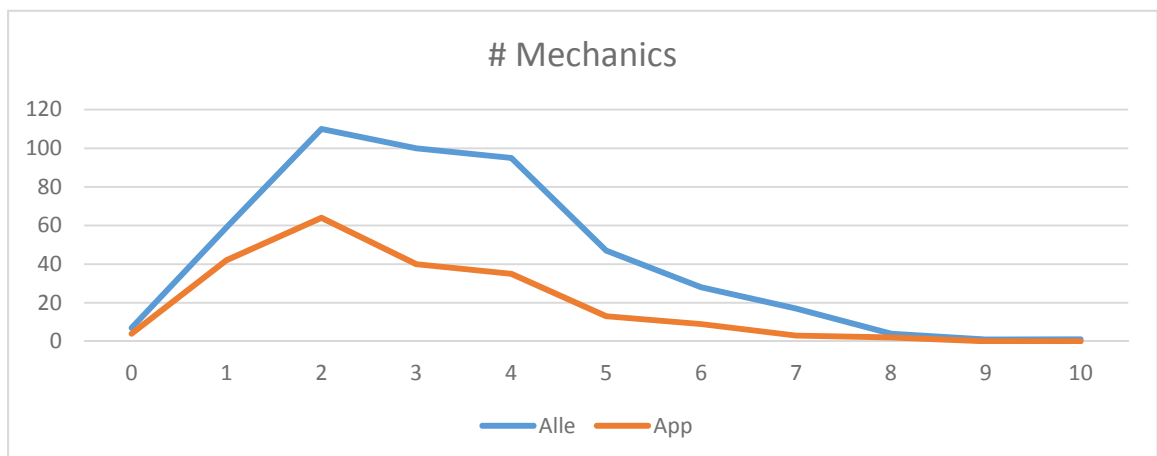


**Bild 8: Shop vs. Avg. Rating**

Ob der deutliche Unterschied zwischen dem Avg. Rating und dem Shop Rating eine Relevanz hat, ist zunächst fraglich. Einerseits kann davon ausgegangen werden, dass die Mitglieder von *BGG* kritischer bei der Bewertung der Spiele sind, da sie einen besseren Einblick in den Umfang der Alternativen und vergleichbare Spiele haben. Andererseits werden in den App Stores nicht zwangsläufig die Spiele an sich, sondern die Qualität der digitalen Umsetzungen bewertet. Zusätzlich werden durch Marketingmaßnahmen – z.B. das Erscheinen eines Dialogs in der App, der zur positiven Bewertung auffordert – die Bewertungen verzerrt.

### 3.7 Mechanics

Wie Bild 9 zeigt, haben die meisten Brettspiele zwischen zwei und vier Mechaniken. Oberhalb der vier Mechaniken ist ein rascher Abwärtstrend erkennbar. Ca. 16% der Brettspiele haben fünf oder sechs Mechaniken und gerade mal 4,9% der Brettspiele haben sieben oder mehr Mechaniken. Bei den digitalen Umsetzungen ist zu erkennen, dass die Mehrheit der Spiele zwei Mechaniken besitzen. Genauer, haben über 70% aller digitalen Umsetzungen drei oder weniger Mechaniken.



**Bild 9: # Mechanics**

Bei der Anzahl der Mechaniken kommt es vorrangig auf die einzelne Mechanik und ihrem Beitrag zur Komplexität/Kompliziertheit des Spieles an. So ist z.B. die Mechanik *Dice Rolling* – Würfeln – weniger kompliziert als die Mechanik *Memory* – das Merken und Wiedergeben von Elementen, Informationen oder Spielabläufen.

Dass die Mehrheit der digitalen Umsetzungen zwischen ein und drei Mechaniken liegt zum einen daran, dass bereits bei dem analogen Pendant nicht sinnlos Mechaniken eingesetzt werden, da dies das Brettspiel nicht zwangsläufig besser machen würde. Andererseits bedeuten mehr Mechaniken auch immer, dass die Komplexität/Kompliziertheit des Spiels gesteigert und damit das Entwickeln der digitalen Umsetzung schwieriger wird. Weiterhin unterstreicht die Bevorzugung weniger Mechaniken den Ansatz der *Casual Games* (Manzius 2012).

Mehr als zwei oder drei Mechaniken erklären sich auch dadurch, dass einige Kombinationen zwangsläufig sind. Z.B. gehen die Mechanik *Hand Management* – Punkte für das Ausspielen von Karten z.B. in einer bestimmten Reihenfolge – und die Mechanik *Card Drafting* – das Ziehen von Karten – oft einher.

#### 4 Limitationen und Ausblick

Im Rahmen dieser ersten Analyse wurde gezeigt, dass mit einfachen Mitteln bereits erste mögliche Interpretationen über Potentiale zur Digitalisierung getroffen werden können. Jedoch wurden verschiedene Faktoren zunächst nicht berücksichtigt, weswegen diese Aussagen noch beschränkt sind.

Der Analyse liegt der Gedanke der explorativen Statistik „Was ist an einer Verteilung eines Merkmals bemerkenswert [...]?“ (Polasek 1994) zugrunde. Allerdings wurde im Rahmen dieser ersten Analyse ausschließlich Wert auf eine univariate deskriptive Statistik gelegt. Dies muss in weiteren Schritten einerseits auf eine bi- und multivariate Analyse erweitert werden, um mögliche Hypothesen von vornherein auszuschließen bzw. weitere zu generieren. Andererseits müssen mittels Verfahren der induktiven Statistik die erzeugten Hypothesen gesichert werden.

Eine Notwendigkeit für die multivariate Analyse zeigt sich z.B. anhand der Merkmale *Playing Time*, *Max. # Players* und *Min. # Players*. So stellt sich die Frage, ob die Spielzeit durch die Anzahl der Spieler – z.B. durch Warten o.ä. Verzögerungen – bedingt ist. Eine KI könnte diese Zeit drastisch verkürzen, wodurch ein Versatz der Spielzeit auf dem mobilen Endgerät stattfinden würde. Sämtliche digitale Umsetzungen wurden mittels des zuvor beschriebenen Verfahrens ausfindig gemacht. Dies schließt nicht aus, dass bestehende oder nicht mehr verfügbare digitale Umsetzungen nicht gefunden wurden. Auch konnten sich in Planung oder Entwicklung befindliche digitale Umsetzungen nicht mit einbezogen werden. Daher sollen zur Vervollständigung der Daten die Spieleverlage befragt werden.

Weiterhin steht der Umsetzung nicht zwangsläufig eine technische Hürde im Weg. So kann die fehlende Umsetzung auch an etwaigen Lizenzvereinbarungen, fehlendem Interesse oder Weitblick des Herausgebers, sowie der Wahl der falschen Partner liegen. Somit ist nicht nur eine Befragung der Spieleverlage, sondern auch der an der digitalen Umsetzung beteiligten Partner relevant.

In den nächsten Schritten sollen zunächst die Analyse auf die multivariate Betrachtung erweitert und die gefundenen Hypothesen dadurch ergänzt sowie durch induktive Verfahren geprüft werden. Parallel wird eine Befragung der Spieleverlage und deren Partner intendiert. Ziel ist es, die gewonnenen Erkenntnisse möglichst zu abstrahieren und deren Gültigkeit in anderen Bereichen zu



prüfen, um ein allgemeingültiges Mittel für Wirtschaft und Forschung zur Beurteilung des Potentials zur Digitalisierung in dieser Branche zu schaffen. Letztlich könnte angestrebt werden, anhand eines Kriterienkatalogs das Umsetzungspotential eines analogen Brettspiels in eine digitale App zu ermitteln.

## 5 Literatur

- App Annie (o.J.) App Annie. <https://www.appannie.com/>. Abgerufen am 25.10.2015
- Basole RC, Karla J (2012) Value Transformation in the Mobile Service Ecosystem: A Study of App Store Emergence and Growth. *INFORMS Service Science* 4(1):24–41
- BGG (o.J. a) BoardGameGeek. <http://boardgamegeek.com>. Abgerufen am 27.10.2015
- BGG (o.J. b) Guide to BoardGameGeek. [http://boardgamegeek.com/wiki/page/Guide\\_to\\_BoardGameGeek](http://boardgamegeek.com/wiki/page/Guide_to_BoardGameGeek). Abgerufen am 27.10.2015
- BGG (o.J. c) BoardGameGeek History. [http://boardgamegeek.com/wiki/page/BoardGameGeek\\_History](http://boardgamegeek.com/wiki/page/BoardGameGeek_History). Abgerufen am 27.10.2015
- BGG (o.J. d) Welcome to BoardGameGeek. [http://boardgamegeek.com/wiki/page/Welcome\\_to\\_BoardGameGeek](http://boardgamegeek.com/wiki/page/Welcome_to_BoardGameGeek). Abgerufen am 27.10.2015
- BGG (o.J. e) XML API. [https://boardgamegeek.com/wiki/page/BGG\\_XML\\_API](https://boardgamegeek.com/wiki/page/BGG_XML_API). Abgerufen am 08.07.2015
- BGG (o.J. f) XML API v2. [https://boardgamegeek.com/wiki/page/BGG\\_XML\\_API2](https://boardgamegeek.com/wiki/page/BGG_XML_API2). Abgerufen am 08.07.2015
- BITKOM (2014) Pressekonferenz – Studie >>Kinder und Jugend 3.0<<. [https://www.bitkom.org/Presse/Anhaenge-an-PIs/2014/April/BITKOM\\_PK\\_Kinder\\_und\\_Jugend\\_3\\_0.pdf](https://www.bitkom.org/Presse/Anhaenge-an-PIs/2014/April/BITKOM_PK_Kinder_und_Jugend_3_0.pdf). Abgerufen am 31.10.2015
- blendle (o.J.) blendle. <https://blendle.com/signup/kiosk>. Abgerufen am 30.10.2015
- BMBF (o.J.) Digitale Wirtschaft und Gesellschaft. <https://www.bmbf.de/de/digitale-wirtschaft-und-gesellschaft-148.html>. Abgerufen am 30.10.2015
- Buck C, Dettweiler C, Eymann T (2014) Informationsökonomische Einordnung mobile Applikationen. *HDM Praxis der Wirtschaftsinformatik* 51(2):188-198
- Grubb J (2015) Kid’s gaming makes up nearly 8% of mobile game spending worldwide. <http://venturebeat.com/2015/08/18/kids-gaming-makes-up-nearly-8-of-mobile-game-spending-worldwide/>. Abgerufen am 31.10.2015
- Hanelt A, Piccinini E, Gregory RW, Hildebrandt B, Kolbe LM (2015) Digital Transformation of Primarily Physical Industries – Exploring the Impact of Digital Trends on Business Models of Automobile Manufacturers. In: Thomas O, Teuteberg F (Hrsg) *Proceedings der 12. Internationalen Tagung Wirtschaftsinformatik (WI 2015)*:1313-1327. Osnabrück
- ICv2 (2015) HOBBY GAMES MARKET CLIMBS TO \$880 MILLION. According to New ICv2 Report. <http://icv2.com/articles/markets/view/32102/hobby-games-market-climbs-880-million>. Abgerufen am 30.10.2015

- IDC (o.J.) Smartphone OS Market Share, 2015 Q2. <http://www.idc.com/prodserv/smartphone-os-market-share.jsp>. Abgerufen am 30.10.2015
- LinkedIn (o.J.) App Annie. <https://www.linkedin.com/company/app-annie>. Abgerufen am 30.10.2015
- Manzius C (2012) Casual Games auf dem Vormarsch. Media Spectrum 4-5:32-33
- Müller MU (2015) Digitalkiosk Blendle: Schlechter Text? Dann gibt's das Geld zurück. SPIEGEL ONLINE. <http://www.spiegel.de/netzwelt/web/blendle-digitalkiosk-startet-in-deutschland-partner-bekannt-a-1037837.html>. Abgerufen am 30.10.2015
- n-tv (2014) Die Deutschen und ihre Brettspiele. Neues „Spiel des Jahres“ führt in den Orient. <http://www.n-tv.de/panorama/Neues-Spiel-des-Jahres-fuehrt-in-den-Orient-article13211741.html>. Abgerufen am 30.10.2015
- Piccinini E, Gregory RW, Kolbe LM (2015) Changes in the Producer-Consumer Relationship - Towards Digital Transformation. In: Thomas O, Teuteberg F (Hrsg) Proceedings der 12. Internationalen Tagung Wirtschaftsinformatik (WI 2015):1634-1648. Osnabrück
- Polasek W (1994) EDA Explorative Datenanalyse: Einführung in die deskriptive Statistik. 2. Auflage. Springer, Berlin Heidelberg
- Schäfer S (2002) Konsequenzen von Digitalisierung und Internet auf die Printmedien. Fachjournalist 4:8
- Spielbar (o.J.) Alterskennzeichen. <http://www.spielbar.de/neu/praxiswissen-computerspiele/jugendschutz/alterskennzeichen/>. Abgerufen am 30.10.2015

# Privacy with Secondary Use of Personal Information

Sven Wohlgemuth<sup>1</sup>, Kazuo Takaragi<sup>2</sup>, and Isao Echizen<sup>3</sup>

<sup>1</sup> Visiting Researcher, Frankfurt/Main, wohlgemuth@acm.org

<sup>2</sup> National Institute of Advanced Industrial Science and Technology, Tokyo,  
kazuo.takaragi@aist.go.jp

<sup>3</sup> National Institute of Informatics, Tokyo, iechizen@nii.ac.jp

## Abstract

Secondary use of personal information is of essential importance for the Internet of Things. The main application is resilience. Biometrics is an example for support of resilience in times of a natural disaster. The primary use of biometrics is to identify people; a secondary use is to improve healthcare services for affected people. This requires information sharing with third parties. The challenge faced for reliable support of the Internet of Things is safety. Special cases of security systems achieve safety for information flow, but they don't scale for secondary use. Their users lose control on their identity. With the aim of improving usability of security, this research-in-progress proposes a multilateral information flow control. This is privacy as understood with informational self-determination. The key is usage control with secure delegation of rights and a secondary use of personal security-related information as Open Data.

## 1 Resilience and Safety

Reliable support of resilience is the goal of data-centric services of the Internet of Things. The common understanding of resilience is to prevent a system from failure by taking its dependencies into account for a predictive risk management. In spite of incidents of any kind, state transitions of the affected system should result in a state of equilibrium, if not even improve the previous one. The support of Internet of Things is to provide information on incidents in accordance to a given service based on context data of the physical environment and their aggregation with existing information in real-time. Such a search engine is a data-centric service and personalized in accordance to its users. Its reliability depends on authentic data. The higher the quantity of authentic data is the lower is the error of the derived information. An example is biometrics (Yamada 2015).

At the Great East Japan Earthquake and the following tsunami on March 11, 2011, information and communication systems have been severely damaged. Information about the identity of the refugees including information about their medical history was not sufficiently available. This has threatened safety of medical treatment. However, the national academic Internet backbone Science Information NETwork – SINET was with its services always available in the disaster area (Urushidani and Aoki 2011). As a lesson learned, a cloud computing information infrastructure for

improving safety in real-time is in development. Its primary use of biometrics is authentication in an application elsewhere, e.g., a secondary use for supporting on-site emergency forces in their medical treatment, may be considered (JAISA 2015).

Reliability of data-centric services is inevitably threatened by misuse of information. For instance, wolves attack on biometric authentication can achieve a false acceptance rate of 60% (Ohki and Otsuka 2014). The origin of a misuse doesn't have to be malicious. If general access control is used, a statement whether a certain access on information will take place in the future is undecidable. This is the safety problem of access control (Harrison et al. 1976). Even though safety is decidable for special cases of access control systems (Sandhu 1993), its use for information sharing results paradoxically for its users in losing control on their identity.

With the aim of improving usability of security in that users control their identity, this research-in-progress article contributes with a scalable information security infrastructure for secondary use. Following Shannon's control by transparency (Shannon 1948), the difference to access control is secure delegation of rights as obligations on the usage of information with acceptable Byzantine Agreement on their observations. The consensus on these digital evidences become available as Open Data. Information on safety is derived by a privacy-enhanced data-centric AAA(A) service.

Section 2 shows that the use of (anonymized) information for statistics needs to be protected as for personal information. Section 3 presents our proposal for control by transparency with a privacy-enhanced AAA(A) service. Section 4 introduces to specifying authorized information sharing with secure delegation of rights. Section 5 describes their use for a prediction of privacy for secondary use. Section 6 shows that related work neither discuss safety nor address secondary use. With an open call for contributions in section 7, we look forward to contribute to a sustainable society.

## 2 Safety and Compliance

Compliance aims at protecting safety and security of users and investors while supporting free flow of information (Senate and House of Representatives of the United States of America in Congress assembled 2002, European Commission 1995). Recently, legal regulations on information processing for business models stipulate resilience (European Commission 2009). They specify minimal security principles to achieve safety of authorized information flows by their isolation. Beside the identities of the participants of an information processing, additional parameters such as time and purpose of the information processing must be taken into account for enforcing isolation. In case of personal data, erasure of personal information is an additional requirement.

For enforcing compliance, regulations stipulate control by transparency. Personal liability of the data consumer holds for the authenticity of the report. The instruments for enforcement are an adequate risk management with internal security controls and an audit by an independent auditor. However, the regulations rule out anonymized information and secondary use with statistics and for cases of severe interests of the public, e.g., critical infrastructure protection. Paradoxically, these exceptional cases threaten compliance and so resilience due to the safety problem of access control.

With type-safety (Sandhu 1993) a statement on future access is decidable under certain conditions, which unfortunately don't fit to compliance of secondary use. It assumes access control commands with at most three parameters and a monotonic, acyclic typed access control model. The first condition contradicts with the requirement of an information exchange with third parties on supporting at least four parameters out of: data provider, data consumer,  $data_1, \dots, data_n$ , purpose,

and time. The second means that irreversible revocation, i.e., erasure of information, is not satisfiable. Aggregation of information at a data-centric services, in turn, increases information asymmetry in contrary to users. This, in turn, is a hurdle for an information market (Laudon 1996).

Whenever integrity or confidentiality, respectively, is more important, the revised Bell-LaPadula BLP or Biba model, respectively, achieves isolation of information flow (Sandhu 1993). Ordered clearance classes for write up/read down or vice versa achieve safety. However, the lattice-based access control model doesn't allow several security classifications of the same subject. This allows information flow only within the same security class. For a secondary use, the data-centric service as a third party must belong for its read and write operations to the same security class of its users. If integrity only should be enforced, the data-centric service must know all information and writes it down to users, or the information must be unclassified. The Chinese-Wall security model achieves for primary use confidentiality, integrity, and availability; the latter with a disclosure of anonymized information. For a secondary use, the third party must know all information on safety violations.

Safety of decentralized access control requires a Byzantine Agreement on distributed security-related information. This, in turn, requires a reliable PKI and running the consensus protocol in the synchronous time model (Pfitzmann and Waidner 1992). Latter impedes availability of information for improving replies to any search request as required for resilience. Relaxing compliance in exceptional cases for support of a trusted third party is not an option. Errors on classification and, hence, on derived information are an integrity breach about the users' information processing (Jernigan and Mistree 2007). This holds also for the use of anonymized information for statistics, since their aggregation enables the algorithm to re-identify the user (de Montjoye et al. 2015).

### **3 Control by Privacy-Enhanced AAA(A)**

Obligations extend access control to usage control with liveness properties for safety (Park and Sandhu 2004). Their formalization for enforcement shows that not all obligations are observable but can become so during run-time (Pretschner et al. 2006). Similar to access control, enforcement of usage control requires a Byzantine Agreement on observation of obligations as digital evidences. Hence, digital evidences are Open Data. Not all observations on accountability of information processing are decidable (Hamlen et. al 2006), but compensation complements observation of obligations. Compensations are enforceable actions, which can take place before access has been granted and after an incident and its accountability has been detected.

Enforcement of usage control extends the standardized IT security infrastructure of the Internet is the AAA Authorization Framework (Vollbrecht et al. 2000). An AAA service enforces an XACML access control policy. Due to vulnerabilities for compliance by the inevitable safety problem, Accountability on information is required as the additional fourth (A) to observe obligations. In contrary to access control, this public AAA(A) service achieves availability of an acceptable Byzantine Agreement on safety, if weakening the requirements on consistency or partition-tolerance (Gilbert and Lynch 2002). Since all users are vulnerable, the threat model of (Dolev and Yao 1983) holds: Enforcement with the AAA(A) service is based on secure multi-party cryptographic protocols, the relations of a cryptographic key to attributes such as the master identity of the given user are authentic, and information exchange is not restricted. In addition, the AAA(A) service is itself a public directory of its users' classification as subjects, their information as objects, and authorizations of usage control. This has two consequences: (1) Byzantine Agreement has to be solvable with cryptographic means without restricting availability of digital evidences; and (2)

the users and their information must be unobservable towards the AAA(A) service to impede information leakage while being at the same time accountable.

Requirement (1) is realized in practice by the block chain as a public directory (Nakamoto 2008). Several economic applications for digital coin based exchange systems show its acceptance in practice (Zohar 2015). A block chain realizes a statistically trusted third party by self-organization of vulnerable users, but is vulnerable to partition-tolerance. Consistency can be tolerated as long as the risk threshold of the corresponding user is compensated. Requirement (2) is Multilateral Security, which aims at a balance between these opposite security interests (Rannenberget al. 1999). In the application of biometrics, privacy-enhanced wearables are additionally required to impede non-authorized information collection and impersonation (Yamada et al. 2013). Adapting Multilateral Security on secondary use results in privacy as informational self-determination.

Our design of such a privacy-enhanced AAA(A) service consists of three block chains: one for each (a) Privacy Authorization, (b) Privacy Accountability, and (c) Privacy Accounting in documenting disclosure of derived (personal) information. Contents of the block chains are either (a) delegated and revoked obligations, (b) evidences on actual information flow, and (c) statements on safety of information processing of users and IT systems, hence, on their reputation. A block represents state transitions as an information exchange between a data provider and data consumer. As a compensation, a data consumer “pays” a financial value to the data provider, e.g., coin or profit sharing on the further processing of this information. In return, the data consumer gets a verifiable statement on the authentication of this data from the data provider and the block chains. State transitions are proven and documented by a user in the successful role of a miner following the proof-of-work of the block chain. A successful miner takes the role of a certifier and is therefore liable for the proof-of-work. As compensation for taking this risk of liability, the successful miner gets reward for the new block and a transaction fee from the data provider. An attested state transition takes for Bitcoin in average 10 minutes (Nakamoto 2008).

## 4 Secure Delegation of Rights

Secure delegation of rights specifies isolation of information flows for privacy in the Internet of Things (Sonehara et al. 2011). It considers users’ liability for compliance in that users control access on their information. In following trust management in that access request in distributed information systems is decided on the authenticity of an attribute-based credential, secure delegation of rights consists of two multi-party cryptographic protocols for authorization with credentials. These protocols on authorization and revocation of authorization extends trust management with the secure delegation of rights. Secure means safety in that users can decide case-by-case on authorizing data consumers for a disclosure of users’ information to a third party, without losing control on using their information at the time of the authorization.

Cryptographic credential systems in general, however, don’t support secondary use. If using them for authorization of a secondary use, users lose control on their identity due to the all-or-nothing property of these identity management systems. Secure delegation of rights enhances trust management, in that binding to credentials of the users to the main cryptographic key remains while authorizing data consumers with certain obligations on using certain information for a certain purpose. The same holds for the revocation protocol for stopping certain usage of information and request for erasing it. Both protocols for secure delegation of rights require a digital directory at a

third party. While using the block chain-based AAA(A) service, these attested obligations become public, whereas the participants are represented by a pseudonym.

## 5 Predicting Privacy for Secondary Use

Enforcement of privacy for secondary use requires digital evidences at least on an information exchange in addition to those on the used (biometric) authentication terminal. A predictive audit with  $k$ -anonymized digital evidences gives a user statistical control on confidentiality (Accorsi et al. 2008). We extend this approach for audit on confidentiality and authentication of information. Depending on the role of a user in data provider and data consumer, an audit predicts privacy before authorization is granted to a data consumer or incoming information from a data provider is processed. Secure delegation of rights of a user spans a directed knowledge graph on expected exchanges of the user's information. Whenever a user receives information from others or is about to delegate rights, he gets a statement on safety by spanning the tree of authorized and the one of actual information exchanges and compare these two trees to find a contradiction to the policy.

Since delegated rights are cryptographic credentials, the check on authentication is the same as for evaluating a PKI from a user's view (Wohlgemuth 2015). An audit on confidentiality requires digital evidences on the actual information flow. This is the purpose of a distributed data provenance system for information accountability. Our system consists of two multi-party cryptographic protocols for tagging current information exchanges and for their verification. The tagging protocol spans the directed graph of the current information flow between identities. These cryptographic evidences are documented in a public digital directory: the block chain-based AAA(A) service. The verification protocol reads evidences from the block chain and checks the relation between this graph with the authorizations of the identities.

This audit service is itself a data-centric service. It scales with the information of digital evidences due to its mathematical abstraction for interoperability on different PKI. In order to improve its statistical statement as reply on a search query on safety, additional digital evidences on secondary use will be added. Improvements of a user's query in its unambiguity should be achieved by aggregation of different sources as for the semantic web (Hahn et al. 2010).

## 6 Related Work

Related work on solving the safety problem of information flow control considers mainly primary use and encryption for a secondary use. Privacy of disclosed personal information is considered with access control and the assumption of a trusted third party. The threat model considers its users as malicious (Maffei et al. 2015). This is contrary to resilience and compliance as shown. Homomorphic encryption allows an exchange of information for joint computation on biometrics (Karvelas et al. 2014), however not for secondary use. A latest agenda for security and privacy in industrial Internet of Things proposes access control (Sadeghi et al. 2015). It neither addresses the purpose of Internet of Things for value-added services by a secondary use nor considers the safety problem. Information accountability considers that access control and encryption are not suitable for protecting privacy (Weitzner et al. 2008). The focus with control by transparency is on controlling the use of data for primary use. Even though, this work proposes a data provenance system to protect privacy, it doesn't consider safety for secondary use.

## 7 Resilient Infrastructure for Value-Added Chains

With the increasing dependency of social infrastructures on reliable support of Internet of Things, security becomes a public asset. At present, improvement of security for resilience with data-centric services depends on those, who control access on the required security-related information. Our proposed extension for the current trust infrastructure follows Privacy by Design. We hope that we contribute to prosperous business continuity of also small and medium enterprises in secure delegating enforcement of privacy to third parties while capital, labor, and ownership of information are still essential instead of being dependent on the centralized control on information. With an open call for contributions to a resilient infrastructure for valued-added services, we look forward to a constructive cooperation for creating a sustainable society.

## 8 References

- Accorsi R, Sato Y, Kai S (2008) Compliance Monitor for Early Warning Risk Determination. *WIRTSCHAFTSINFORMATIK* 50(5):375—382
- Dolev D, Yao, AC (1983) On the Security of Public Key Protocols. *IEEE Transactions on Information Theory* 29(2):350—357
- European Commission (1995) Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data. *OJ L281 (395L0046)*:31—50
- European Commission (2009) Directive 2009/140/EC of the European Parliament and of the Council of 25 November 2009 amending Directives 2002/21/EC on a common regulatory framework for electronic communications networks and services. *OJ L 337*:37—69
- Freire E, Hofheinz D, Kiltz E, Paterson K (2013) Non-Interactive Key Exchange. In: PKC 2013. Nara
- Gilbert S, Lynch N (2002) Brewer's Conjecture and the Feasibility of Consistent, Available, Partition-Tolerant Web Services. *ACM SIGACT News* 33(2):51—59
- Hoang HH, Nguyen TM, Tjoa AM (2009) A Semantic Web Based Approach for Context-Aware User Query Formulation and Information Retrieval. In: Alkhatib G, Rine D (eds) *Web Engineering Advancements and Trends*. Information Science Reference, Hershey NY
- Hamlen KW, Morrisett G, Schneider FB (2006) Computability Classes for Enforcement Mechanisms. *ACM Transactions on Programming Languages and Systems* 28(1):175—205
- Harrison MA, Ruzzo WL, Ullman JD (1976) Protection in Operating Systems. *CACM* 19(8):461—471
- JAISA (2014) Research and development of victim support system using biometric authentication. (in Japanese) <http://www.jaisa.jp/pdfs/150430/v1-1.pdf> Accessed on December 14, 2015
- Jernigan C, Mistree BFT (2007) Gaydar: Facebook friendships expose sexual orientation. *First Monday* 14(10)
- Karvelas NP, Peter A, Katzenbeisser S, Tews E, Hamacher K (2014) Privacy-Preserving Whole Genome Sequence Processing through Proxy-Aided ORAM. In: WPES'14. Scottsdale AZ
- Laudon KC (1996) Markets and Privacy. *CACM* 39(9):92—104



- Maffei M, Malavolta G, Reinert M, Schröder G (2015) Privacy and Access Control for Outsourced Personal Records. In: IEEE Symposium for Security and Privacy 2015, San Jose CA
- de Montjoye YA, Radaelli L, Singh VK, Pentland A (2015) Unique in the shopping mall: On the reidentifiability of credit card metadata. *Science* 347(6221):536—539
- Nakamoto S (2008) Bitcoin A Peer-to-Peer Electronic Cash System. <https://bitcoin.org/bitcoin.pdf>  
Accessed on September 27, 2015
- Ohki T, Otsuka A (2014) Theoretical vulnerability in likelihood-ratio-based biometric verification. In: 2014 IEEE IJCB. Clearwater FL
- Park J, Sandhu RS (2004) The UCON<sub>ABC</sub> Usage Control Model. *ACM Transactions on Information and Systems Security* 7(1):128—174
- Pfitzmann B, Waidner M (1992) Unconditional Byzantine Agreement for any Number of Faulty Processes. In: STACS '92. Paris
- Pretschner A, Hilty M, Basin, D (2006) Distributed usage control. Müller G (ed.) Personalization in privacy-aware highly dynamic systems *CACM* 49(9):39—44
- Rannenber K, Pfitzmann A, Müller G (1999) IT Security and Multilateral Security. In: Müller G, Rannenber K (eds.): *Multilateral Security in Communications – Technology, Infrastructure, Economy*. Addison-Wesley-Longman, München
- Sadeghi AR, Wachsmann C, Waidner M (2015) Security and Privacy Challenges in Industrial Internet of Things. In: DAC '15. San Francisco CA
- Sandhu RS (1993) Lattice-Based Access Control Models. *Computer* 26(11):9—19
- Senate and House of Representatives of the United States of America in Congress assembled (2002) Sarbanes-Oxley Act of 2002. Pub. L. No. 107–204, 116 Stat. 745, H.R. 3763—66
- Shannon CE (1948) A mathematical theory of communication. *The Bell System Technical Journal* 27(3):379—423
- Sonehara N, Echizen I, Wohlgemuth S (2011) Isolation in Cloud Computing and Privacy-Enhancing Technologies. *BISE* 3(3):155—162
- Urushidani S, Aoki M (2011) Design and Implementation of Reliable Multi-layer Service Network. In: 3<sup>rd</sup> International Congress on ICUMT. Budapest
- Vollbrecht J, Calhoun P, Farrell S, Gommans L, Gross G, de Bruijn B, de Laat C, Holdrege M, Spence D (2000) AAA Authorization Framework – RfC 2904. Network Working Group
- Weitzner DJ, Abelson H, Berners-Lee T, Feigenbaum J, Hendler J, Sussman GJ (2008) Information Accountability. *CACM* 51(6): 82—87
- Wohlgemuth S (2015). Resilience by Usable Security. In: *Usable Security and Privacy, Mensch und Computer 2015*. Stuttgart
- Yamada T, Gohshi S, Echizen I (2013) Privacy Visor: Method based on Light Absorbing and Reflecting Properties for Preventing Face Image Detection. In: *IEEE SMC 2013*. Manchester
- Yamada A (2015) A Generalization of ISO/IEC 24761 to Enhance Remote Authentication with Trusted Product at Claimant. In: Federrath H, Gollmann D (eds.) *IFIP SEC 2015*. Hamburg
- Zohar A (2015) Bitcoin: Under the Hood. *CACM* 58(9):104—113