



TECHNISCHE UNIVERSITÄT
CHEMNITZ

TU Chemnitz
Fakultät für Wirtschaftswissenschaften
Professur für Wirtschaftsinformatik – Geschäftsprozess- und
Informationsmanagement
Prof. Dr. Barbara Dinter

Tagungsband zum
**20. Interuniversitären Doktorandenseminar
Wirtschaftsinformatik**

(Hrsg.)
Barbara Dinter,
Lisa Frenzel,
Peter Gluchowski

Chemnitz, den 27.10.2016

Tagungsband zum
20. Interuniversitären Doktorandenseminar Wirtschaftsinformatik

Herausgeber: Barbara Dinter, Lisa Frenzel, Peter Gluchowski
Verlag: Universitätsverlag der Technischen Universität Chemnitz
Erscheinungsjahr: 2017
ISBN: 978-3-96100-009-8
URN: <http://nbn-resolving.de/urn:nbn:de:bsz:ch1-qucosa-216683>

Der Tagungsband zum 20. Interuniversitären Doktorandenseminar Wirtschaftsinformatik dient der Darstellung des jeweils aktuellen Standes der teilnehmenden Promovierenden zu ihren Promotionsvorhaben. Die Autoren sind deshalb i.d.R. dankbar für kritische Hinweise. Die individuellen Kontaktdetails der Autoren am Ende des jeweiligen Beitrages sollen dafür genutzt werden.

Vorwort

Das Interuniversitäre Doktorandenseminar Wirtschaftsinformatik ist eine regelmäßige Veranstaltung, in deren Rahmen Doktoranden der Universitäten Chemnitz, Dresden, Freiberg, Halle, Ilmenau, Jena und Leipzig ihr Promotionsprojekt präsentieren und sich den kritischen Fragen der anwesenden Professoren und Doktoranden aller beteiligten Universitäten stellen. Auf diese Weise erhalten die Promovierenden wertvolles Feedback zu Vorgehen, Methodik und inhaltlichen Aspekten ihrer Arbeit, welches sie für ihre Promotion nutzen können.

Darüber hinaus bietet das Interuniversitäre Doktorandenseminar Wirtschaftsinformatik eine Plattform für eine fachliche Auseinandersetzung mit aktuellen Themen und sich ankündigenden Trends in der Forschung der Wirtschaftsinformatik. Zudem wird ein akademischer Diskurs über die Grenzen der jeweils eigenen Schwerpunkte der Professur hinaus ermöglicht.

Das nunmehr 20. Jubiläum des Doktorandenseminars fand in Chemnitz statt. Der daraus entstandene Tagungsband enthält fünf ausgewählte Beiträge zu den Themenfeldern Service Engineering, Cloud-Computing, Geschäftsprozessmanagement, Requirements Engineering, Analytics und Datenqualität und zeigt damit anschaulich die Aktualität und Relevanz, aber auch die thematische Breite der gegenwärtigen Forschung im Bereich Wirtschaftsinformatik.

Chemnitz, Oktober 2016

Barbara Dinter,

Lisa Frenzel,

Peter Gluchowski

Vorgestellte Beiträge

Building Blocks and Granularity of Logistics Services <i>Michael Glöckner</i>	6
Privacy-Erhaltung in kollaborativen Geschäftsprozessen in der Cloud <i>Björn Schwarzbach</i>	18
Gewichtung von Anforderungen an ein Process Performance Measurement System zur Unterstützung einer ganzheitlichen Working-Capital-Steuerung <i>Marcus Pfitzner</i>	30
Promotionsvorhaben zur Erweiterung des Algorithmus CLTree <i>Martin Thrum</i>	41
Unternehmensübergreifende Stammdatenqualität – Entwicklung eines Hilfsmittels zur Vereinbarung der Qualität für Stammdaten zwischen Unternehmen <i>Thomas Schäffer</i>	49

Building Blocks and Granularity of Logistics Services

Michael Glöckner

Endowed Chair of Logistics Information Systems,
Information Systems Institute, Leipzig University

Abstract

Future logistics is facing serious challenges in terms of concentration on core competencies and digitalization. Objective of this is a higher flexibility of logistics networks and supply chains. Logistics service providers get both the chance and obligation to become more collaborative. The paradigm of cloud logistics is a promising approach to consider both digitalization as well as outsourcing and collaboration. Cloud logistics adopts the basic principles of cloud computing, such as virtualized resources, and encapsulating services from different providers. Precondition for a successful transformation to cloud logistics are (1) an approach for generic building blocks that are able to encapsulate logistics functions and capabilities in terms of real and virtual aspects. Further (2) service granularity is an important issue in terms of modeling existing logistics service portfolios with the generic building blocks. Leading method is the framework of design oriented information systems research.

Keywords

Cloud Logistics, Service Blueprinting, Modularization, Service Granularity Framework

1 Introduction

Future logistics is facing serious challenges in terms of concentration on core competencies (Langley & Long, 2016; Wilding & Juriado, 2004) and digitalization (Kagermann, 2015; Lai, Wong, & Cheng, 2010; Yang, Pan, & Ballot, 2016). Both aim at a higher flexibility of logistics networks and supply chains (Solakivi, Töyli, & Ojala, 2013). Further logistics service providers (LSP) get the chance and obligation to become more collaborative, which improves companies' performance (Olorunniwo & Li, 2010). A promising approach to solve both problems is the paradigm of cloud logistics (Delfmann & Jaekel, 2012; Li, Zhong, Wang, & Cao, 2013). Cloud logistics is built upon the adoption of the basic principles of cloud

computing (Mell & Grance, 2011; Vaquero, Rodero-Merino, Caceres, & Lindner, 2008), such as virtualized resources, interconnection ability of services from different providers and ad-hoc reconfiguration of service bundles. Precondition for a successful translation to cloud logistics are twofold. Generic building blocks that are able to encapsulate logistics functions and capabilities in terms of real and virtual aspects of logistics services are an important aspect in order to enable modular cloud logistics (Delfmann & Jaekel, 2012). Further, service granularity is an important issue in terms of decomposing and modeling existing logistics service portfolios with the generic building blocks (Steghuis, 2006). The paper's goal is to develop a framework of modular building blocks of logistics services that are used to model a distinct granularity level of services in logistics networks. Leading question is: "*How can logistics services be modeled in order to create generic building blocks on a suitable granularity level in logistics networks?*" It results in the particular research questions:

1. What is a suitable level of service granularity for collaboration in logistics networks?
2. What is an appropriate concept of logistics service building blocks?

Following the framework of design oriented information systems research (Hevner, March, Park, & Ram, 2004; Österle et al., 2010), several research methods have been integrated in the certain phases to develop the artifacts. Those methods comprise: systematic literature review (Vom Brocke et al., 2015), extended service blueprinting (Hara, Arai, Shimomura, & Sakao, 2009), domain engineering (Czarnecki & Eisenecker, 2000) and general morphological analysis (Ritchey, 2013). The developed artifacts are a service granularity framework and a conceptual model of logistics services in order to create generic modular cloud logistics service blueprints. With the developed artifacts, the paradigm of cloud logistics is enabled in a conceptual way. Researchers find starting points for extended research on cloud logistics, practitioners find initial ideas of the paradigm in order to implement them into practice.

After the introduction, requirements are presented in the second section. Further, the conceptual models are outlined as well as the influencing literature. The third section merges the both developed artifacts and shows how the concepts enable cloud logistics. The last section concludes the paper, discusses finding and limitations as well as future research directions.

2 Conceptual Models

This section outlines requirements of modular logistics services and presents the developed artifacts, the used methods, and the influencing concepts of literature.

2.1 Requirements

Logistics is a typical service-oriented industry. Its main purpose is the timely and spatial transformation of goods and information (Gudehus & Kotzab, 2012). With an increasing outsourcing and concentration on core competencies (Langley & Long, 2016), LSP are required to *collaborate* with each other in order to increase performance (Olorunniwo & Li,

2010), to handle the dynamic character of logistics networks (Solakivi et al., 2013) and to fulfill complex customer demands by bundling their competencies. This further implies the *integration of external resources and information* of other LSP and customers (Hipp & Grupp, 2005). Common *communication standards* are necessary to enable effective collaboration, service visibility and an increased value cocreation (Rai, Pavlou, Im, & Du, 2012). Within the network different abstraction levels are to be covered and aligned with each other. An *alignment* has to be created from a portfolio of all services in the network on top (Kohlborn, Fieft, Korthaus, & Rosemann, 2009) down to basic logistics functions of the participating LSP. When talking about information sharing and collaboration, *privacy* concerns are of great importance for LSP (Liu, Li, & Zhao, 2009), e.g. because of increased economic competition. Thus, a distinction between information for internal and external use is to be regarded in cross-organizational business collaboration, e.g. see (Norta et al., 2015). Besides general collaboration requirements, the specialties of cloud logistics have to be taken into account. Physical and non-physical logistics resources need to be *virtualized* in order to connect them easily with resources from other LSP in the network or the supply chain (Li et al., 2013).

2.2 Service Granularity Framework

In order to give guidance in the creation of modular logistics services, service granularity has to be conceptualized. Existing literature on service granularity is analyzed and the found concepts are compared. From this, a conceptual framework of service granularity is derived. With this, guidance can be given for the modularization of existing service portfolios. This helps developing services in order to decide if further composition or decomposition is needed for services on particular levels with particular characteristics. A starting point for the conceptualization of service granularity is one of the first definitions of service granularity given by Papazoglou and van Heuvel (2006): ‘service granularity is the scope of functionality exposed by a service’.

Several databases have been taken into account (science direct, web of science, springerlink, emerald) and were searched for the exact phrase “service granularity”. Further, forward and backward search was conducted. In total 139 paper were found and analyzed. The biggest amount (118 paper) comprises the paper of the categories that either were just mentioning service granularity as an ‘important aspect’ of service engineering without explaining it in detail; or did not deal with service granularity at all; or were duplicates.

The second category comprises 10 paper with a detailed description of service granularity. The classical trade-off between fine and coarse grained services is outlined (Kulkarni & Dwivedi, 2008; Steghuis, 2006). Some elementary measurements are presented (Bianchini, Cappiello, Antonellis, & Pernici, 2014; Feuerlicht, 2011; Katzmarzik, 2011; Sindhgatta, Sengupta, & Ponnalagu, 2009). Other paper deal with automatic service identification in terms of decomposition until a certain level of granularity is reached (Chatla et al., 2011; Kim & Doh, 2009; Wang & Li, 2014). Those approaches only deal with quantitative characteristics, but do not distinguish qualitative nor semantic aspects of services. A use case on service granularity measurement is described by Feuerlicht (2007) that cannot be generalized.

Furthermore, the third category comprises 11 paper that describe service granularity very detailed and influence the final conceptualization of the framework. Galster and Bucherer (2008) align business requirements (problem domain) top-down and services (solution domain) bottom-up on a *common level of abstraction*. This mapping level allows a 1:1 connection between both domains. They propose three layers within each domain, similarly to Ma, Zhou, Zhu, and Wang (2009). Erradi, Kulkarni, and Maheshwari (2007) introduce a similar approach, but without distinct levels, in order to match (electronic) services from IT applications (bottom-up) to (business) services from business processes (top-down) on a common granularity level. Further, they emphasize a *canonical schema* and *service meta data management* on the common level in order to facilitate consistent representation as well as syntactic and semantic mapping. *Two orthogonal dimensions of granularity* are inspired by both Huergo et al. (2014) (with a distinction between service hierarchy and service type) as well as Heinrich and Zimmermann (2012) (with a distinction between measures of service width and depth). The hierarchy idea is further supported by Cai, Cui, Shi, Kong, and Yan (2014). Atomic services as well as process, role and goal granularity is differentiated by Liu, He, Wang, Liu, and Li (2015). The trade-off between coarse and fine grained services is discussed (Granell, Díaz, & Gould, 2010; Haesen, Snoeck, Lemahieu, & Poelmans, 2008; Papazoglou & van Heuvel, 2006). In general more coarse grained service appear to be more useful to customers as they contain more functionality and reduce overhead at the same time. Decomposition should be stopped, when further splitting of services would not make sense and companies are encouraged to offer useful service bundles. Interestingly, certain set of pre-defined aspects is very useful to provide guidance during engineering of modular service (Dörbecker & Böhm, 2015).

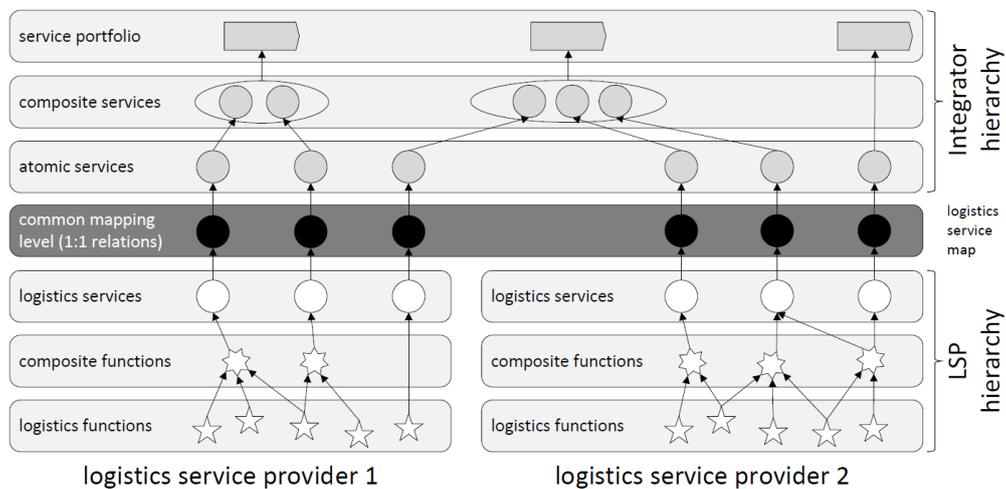


Figure 1: Service granularity framework applied to an example logistics service network.

The resulting service granularity framework, which can be seen in an applied example in Figure 1, consists conceptually of bundles of three provider levels (top, middle, bottom) and a common mapping level between provider level bundles of different hierarchy. The middle

level is used to internally manipulate granularity in terms of service composition and creation of service bundles by one provider within its own hierarchy. The bottom level is used to integrate services and function from service providers of a lower hierarchy. On the contrary the top level is used to present and offer services externally to the next higher hierarchy level in the service network. The objective of the common mapping level is the alignment of differing service descriptions from different providers. It contains virtual entities that represent the commonly shared standard (e.g. description, certain granularity) to which the bottom-level services of the next upward LSP and the top-level services of the next downward LSP must comply. This level is also adequate for the implementation of a canonical schema and the service metadata management. The mapping level appears to be suitable for the allocation of common standards of collaboration, or generic logistics service building blocks, respectively. Hence, the first research question is answered.

2.3 Cloud Logistics Service Blueprint

In order to create building blocks of logistics services for the purpose of cloud logistics, several aspects have to be taken into account. First, an appropriate service engineering approach to create a generic building block has to be identified. Second, general service aspects as well as logistics specific aspects have to be considered. Third, special characteristics of cloud logistics are to be regarded during creation. With this the foundation of a modular system of building blocks on logistics is created. Those ‘lego bricks’ (Papazoglou, 2012; Weißenberg & Springer, 2014) of cloud logistics are strongly inspired by the idea of generic cloud blueprints from the cloud computing domain (Papazoglou, 2012). Through analogical design, the blueprint idea is adapted to the logistics domain in order to create the generic foundation for the lego bricks.

Concerning service engineering methods, the leading method is *extended service blueprinting* of Hara et al. (2009). It is able to describe services that comprise business service aspects as well as electronic service aspects. Further, essential service aspects (Erl, 2009) like interaction between provider and consumer, as well as physical and non-physical input and output can be modeled. Basic distinction is made between the ‘behaviour blueprint’ for hardware-based service provision (= electronic services) and the ‘activity blueprint’ for human-based service provision (= business services). Both are included and connected in the service blueprint. Receiver state parameters (RSP) are used to describe a change of certain aspects of the customer as the goal for customer demand and service provision. The extended service blueprinting approach features the concept of ‘line of interaction’ and the ‘line of visibility’ to model internal and external aspects and enable certain privacy aspects. *Domain engineering* (Czarnecki & Eisenecker, 2000) is used to find common and varying points of a domain in order to integrate the common points into concepts and to model configurable concepts for the varying. Its goal is exactly the creation of re-usable IT artifacts. Finally, the method of *general morphological analysis* (Ritchey, 2013) is used to structure the multidimensional problem complex and to create a morphological field and identify all dimensions of a problem as well as the possible spectrum of common and varying points.

The resulting generic modular building block of logistics services is displayed in Figure 2. Services in general are characterized by a flow of entities, and logistics services in particular rely on the *flows* of goods and information (Gudehus & Kotzab, 2012). Further, a control flow

is necessary to coordinate several logistics objects, modules and services (Weißenberg & Springer, 2014). Services' objective is a certain transformation in order to create a value for the customer. (Erl, 2009; Papazoglou & van Heuvel, 2006). Those *transformation* dimensions could be represented by the seven targets of logistics (Mentzer, Flint, & Kent, 1999). i.e. deliver the right product, to the right location, in the right time, with the right quality, in the right quantity, for the right price, and with the right information. As described by Li et al. (2013), for the purpose of cloud logistics resources need to be virtualized. In this context resources can be basically distinguished into active and passive. Examples for *active resources* that are able to actively conduct dimension transformations (given as examples in parenthesis) are e.g. trucks (location), warehouses (time, quality), picking systems (product, quantity), conveyors (location), sorter (product, quantity), warehouse management systems (WMS) and Transport Management Systems (TMS) (information). Examples for *passive resources* that are used to contain goods or information are, e.g. packing, pallet, trailer, container (according to the physical flow) or e.g. freight documents, pick lists, warehousing contract (according to the informational flow). Domain specific service level agreements (SLA) are part of the varying points as they are case dependent and determined by individual contracts, e.g. lead time, delivery rate, reliability. Different policies are to be included as varying points, as they are always subject of individual contracts. Compliance-related topics like privacy and security are important issues in logistics (Marucheck, Greis, Mena, & Cai, 2011). Further, legal constraints play an important role in supply chains (Li, Zhang, & Li, 2014), e.g. legal regulations on the allowed period of driving and rest in road transport (Regulation on the harmonisation of certain social legislation relating to road transport., 2006), permission to handle dangerous goods (United Nations, 2012), or permission to handle transport entities that are alive. Additionally, non-mandatory regulations can be incorporated, such as environmental, social and/or economic constraints (Anand, Yang, van Duin, & Tavasszy, 2012). Through the standardized input and output flows modules can be connected with each other. The transfer of goods from one module to another is done in real life and has to be confirmed by a human or machine interface to the system. The flow of information comprises data like order number, destination, etc. The control flow starts and ends modules and triggers next module after completion of the former one. Further, important information can be transferred to a planning and monitoring party of the supply chain by the control flow.

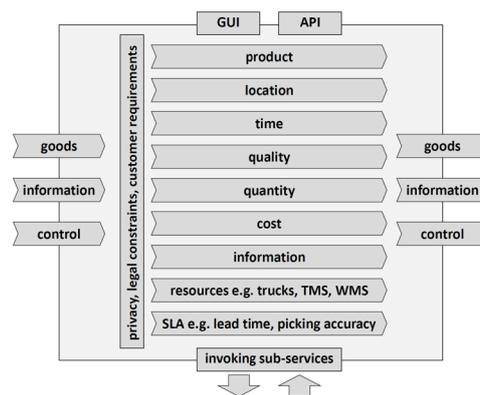


Figure 2: generic service blueprint for cloud logistics.

3 Service Blueprints within the Granularity Framework

The described cloud logistics service blueprints can be used as generic building blocks and a common standard of a network. They comprise general aspects of logistics services and thus can be used as a canonical schema of a logistics network on the common mapping level of the granularity framework. As shown in Figure 3, different LSP can relate their externally offered logistics services (top provider level of the LSP, bottom layer in Figure 3) to the standardized service blueprints (on the common mapping level). Those blueprints are equal to the atomic services of the network. The atomic services comprise all available service in the logistics network. As they are all based on the same generic cloud logistics service blueprint, they are freely inter-connectable. The three standardized interfaces of the several flows (i.e. goods, information, and control) enable their connection. Those services and composite services on the network hierarchy present virtualized models of the services offered by the LSP. Those virtualized services can be used for planning, controlling and monitoring by the integrating party. In the background, an integration from those virtualized services of the integrator's IT systems to the IT systems of the LSP has to be realized by either web interfaces or systems integration. Complex service can be created by composition on the middle level of the network hierarchy. Further, those composite services can be offered to customers of the network. In the example, typical logistics operations happening at an air hub are composed, i.e. truck off-loading, warehousing and subsequent picking concerning given customer specific orders.

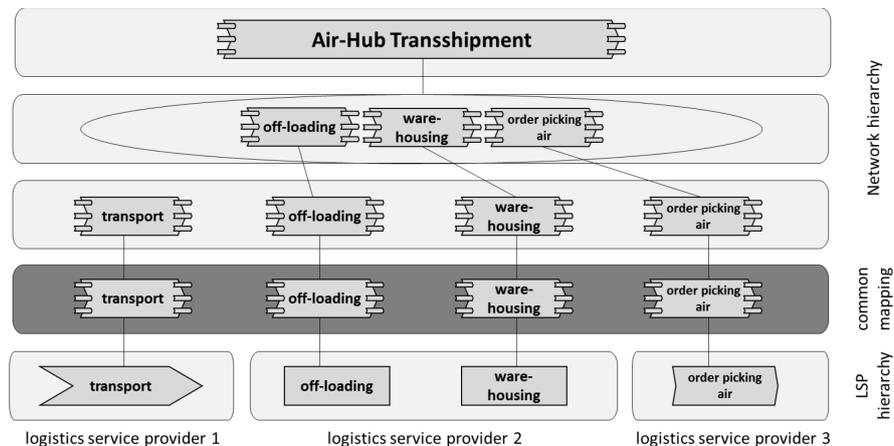


Figure 3: Integration of the cloud logistics service blueprints in the service granularity framework.

4 Conclusion

In this paper two concepts of cloud logistics have been presented and integrated. With the virtualized logistics services and the ability to encapsulate them to more complex composite logistics services the approach of cloud logistics is enabled. By adopting this approach,

networks of LSP could become more flexible because of the facilitated inter-connection and integration with other LSP. Further, by integration via web interfaces even small LSP are enabled to take first steps into a digitalized logistics era.

Limitations of the presented work are a missing evaluation with empirical data. As this paper shows first conceptual steps, a prototypical development could lead to more detailed insights and a demonstrator in order to further include feedback from practicing point of view. Those are next steps of the research.

References

- Anand, N., Yang, M., van Duin, J., & Tavasszy, L. (2012). GenCLOn: An ontology for city logistics. *Expert Systems with Applications*, 39(15), 11944–11960. doi:10.1016/j.eswa.2012.03.068
- Bianchini, D., Cappiello, C., Antonellis, V. de, & Pernici, B. (2014). Service Identification in Interorganizational Process Design. *IEEE Transactions on Services Computing*, 7(2), 265–278. doi:10.1109/TSC.2013.26
- Cai, H., Cui, L., Shi, Y., Kong, L., & Yan, Z. (2014). Multi-tenant Service Composition Based on Granularity Computing. In *2014 IEEE International Conference on Services Computing (SCC)* (pp. 669–676). doi:10.1109/SCC.2014.93
- Chatla, S., Kadam, S., Kolluru, D., Sinha, S., Viswandhuni, A., & Vaidya, A. (2011). Complex networks and SOA: Mathematical modelling of granularity based web service compositions. *Sadhana*, 36(4), 441–461. doi:10.1007/s12046-011-0036-2
- Czarnecki, K., & Eisenecker, U. (2000). *Generative programming: Methods, tools, and applications*. Boston: Addison Wesley.
- Delfmann, W. & Jaekel, F. (2012). *The Cloud - Logistics for the Future?: Discussionpaper*. Retrieved from <http://www.bvl.de/misc/filePush.php?id=18910&name=Discussionpaper+Cloud+Logistic>
- Dörbecker, R., & Böhmman, T. (2015). Tackling the Granularity Problem in Service Modularization. In *Twenty-first Americas Conference in Information Systems (AMCIS)*.
- Erl, T. (2009). *SOA: Principles of service design* (5. print). *The Prentice Hall service-oriented computing series from Thomas Erl*. Upper Saddle River, NJ [u.a.]: Prentice Hall.
- Erradi, A., Kulkarni, N., & Maheshwari, P. (2007). Service Design Process for Reusable Services: Financial Services Case Study. In B. J. Krämer, K.-J. Lin, & P. Narasimhan (Eds.), *Lecture Notes in Computer Science. Service-Oriented Computing – ICSOC 2007* (Vol. 4749, pp. 606–617). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-540-74974-5_56
- Feuerlicht, G. (2007). Service Aggregation Using Relational Operations on Interface Parameters. In D. Georgakopoulos, N. Ritter, B. Benatallah, C. Zircpins, G. Feuerlicht, M. Schoenherr, & H. R. Motahari-Nezhad (Eds.), *Lecture Notes in Computer Science. Service-Oriented Computing ICSOC 2006* (Vol. 4652, pp. 95–103). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-540-75492-3_9
- Feuerlicht, G. (2011). Simple Metric for Assessing Quality of Service Design. In D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, . . . M. Fantinato (Eds.), *Lecture Notes in Computer Science. Service-Oriented Computing* (Vol. 6568, pp. 133–143). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-19394-1_14

- Galster, M., & Bucherer, E. (2008). A Business-Goal-Service-Capability Graph for the Alignment of Requirements and Services. In *2008 IEEE Congress on Services Part 1 (SERVICES-1)* (pp. 399–406). doi:10.1109/SERVICES-1.2008.62
- Granell, C., Díaz, L., & Gould, M. (2010). Service-oriented applications for environmental models: Reusable geospatial services. *Environmental Modelling & Software*, *25*(2), 182–198. doi:10.1016/j.envsoft.2009.08.005
- Gudehus, T., & Kotzab, H. (2012). *Comprehensive Logistics*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Haesen, R., Snoeck, M., Lemahieu, W., & Poelmans, S. (2008). On the Definition of Service Granularity and Its Architectural Impact. In Z. Bellahsene & M. Léonard (Eds.), *Lecture Notes in Computer Science. Advanced Information Systems Engineering* (Vol. 5074, pp. 375–389). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-540-69534-9_29
- Hara, T., Arai, T., Shimomura, Y., & Sakao, T. (2009). Service CAD system to integrate product and human activity for total value. *CIRP Journal of Manufacturing Science and Technology*, *1*(4), 262–271. doi:10.1016/j.cirpj.2009.06.002
- Heinrich, B., & Zimmermann, S. (2012). Granularity metrics for it services. In *Proceedings of the International Conference on Information Systems (ICIS) 2012* (pp. 1–19).
- Hevner, A., March, S., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly*, *28*(1), 75–105.
- Hipp, C., & Grupp, H. (2005). Innovation in the service sector: The demand for service-specific innovation measurement concepts and typologies. *Research Policy*, *34*(4), 517–535. doi:10.1016/j.respol.2005.03.002
- Huergo, R. S., Pires, P. F., Delicato, F. C., Costa, B., Cavalcante, E., & Batista, T. (2014). A systematic survey of service identification methods. *Service Oriented Computing and Applications*, *8*(3), 199–219. doi:10.1007/s11761-014-0161-y
- Kagermann, H. (2015). Change Through Digitization—Value Creation in the Age of Industry 4.0. In H. Albach, H. Meffert, A. Pinkwart, & R. Reichwald (Eds.), *Management of permanent change* (pp. 23–45). New York: Springer Gabler. doi:10.1007/978-3-658-05014-6_2
- Katzmarzik, A. (2011). Product Differentiation for Software-as-a-Service Providers. *Business & Information Systems Engineering*, *3*(1), 19–31. doi:10.1007/s12599-010-0142-4
- Kim, Y., & Doh, K.-G. (2009). Formal Identification of Right-Grained Services for Service-Oriented Modeling. In D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, . . . J. X. Yu (Eds.), *Lecture Notes in Computer Science. Web Information Systems Engineering - WISE 2009* (Vol. 5802, pp. 261–273). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-04409-0_29
- Kohlborn, T., Fieft, E., Korthaus, A., & Rosemann, M. (2009). Towards a Service Portfolio Management Framework. In *ACIS 2009 Proceedings*.
- Kulkarni, N., & Dwivedi, V. (2008). The Role of Service Granularity in a Successful SOA Realization A Case Study. In *2008 IEEE Congress on Services Part 1 (SERVICES-1)* (pp. 423–430). doi:10.1109/SERVICES-1.2008.86
- Lai, K.-h., Wong, C. W., & Cheng, T. (2010). Bundling digitized logistics activities and its performance implications. *Industrial Marketing Management*, *39*(2), 273–286. doi:10.1016/j.indmarman.2008.08.002

- Langley, J., & Long, M. (2016). 2016 Third-Party Logistics Study: The State of Logistics Outsourcing, 20. Retrieved from <http://www.otmbe.org/infotheek/downloads/informatie/652-3pl-report/file>
- Li, C., Zhang, X., & Li, L. (2014). Research on Comparative Analysis of Regional Logistics Information Platform Operation Mode Based on Cloud Computing. *International Journal of Future Generation Communication and Networking*, 7(2), 73–80. doi:10.14257/ijfgcn.2014.7.2.08
- Li, W., Zhong, Y., Wang, X., & Cao, Y. (2013). Resource virtualization and service selection in cloud logistics. *Journal of Network and Computer Applications*, 36(6), 1696–1704. doi:10.1016/j.jnca.2013.02.019
- Liu, C., Li, Q., & Zhao, X. (2009). Challenges and opportunities in collaborative business process management: Overview of recent advances and introduction to the special issue. *Information Systems Frontiers*, 11(3), 201–209. doi:10.1007/s10796-008-9089-0
- Liu, J., He, K., Wang, J., Liu, F., & Li, X. (2015). Service organization and recommendation using multi-granularity approach. *Knowledge-Based Systems*, 73, 181–198. doi:10.1016/j.knosys.2014.10.002
- Ma, Q., Zhou, N., Zhu, Y., & Wang, H. (2009). Evaluating Service Identification with Design Metrics on Business Process Decomposition. In *2009 IEEE International Conference on Services Computing* (pp. 160–167). doi:10.1109/SCC.2009.44
- Maruchek, A., Greis, N., Mena, C., & Cai, L. (2011). Product safety and security in the global supply chain: Issues, challenges and research opportunities. *Journal of Operations Management*, 29(7-8), 707–720. doi:10.1016/j.jom.2011.06.007
- Mell, P., & Grance, T. (2011). The NIST definition of cloud computing. *Computer Security Division, Information Technology Laboratory, National Institute of Standards and Technology Gaithersburg*.
- Mentzer, J. T., Flint, D. J., & Kent, J. L. (1999). Developing a logistics service quality scale. *Journal of Business Logistics*, 20(1), 9.
- Norta, A., Ma, L., Duan, Y., Rull, A., Kõlvart, M., & Taveter, K. (2015). eContractual choreography-language properties towards cross-organizational business collaboration. *Journal of Internet Services and Applications*, 6(1), 1. doi:10.1186/s13174-015-0023-7
- Olorunniwo, F. O., & Li, X. (2010). Information sharing and collaboration practices in reverse logistics. *Supply Chain Management: An International Journal*, 15(6), 454–462. doi:10.1108/13598541011080437
- Österle, H., Becker, J., Frank, U., Hess, T., Karagiannis, D., Krcmar, H., . . . Sinz, E. J. (2010). Memorandum on design-oriented information systems research. *European Journal of Information Systems*, 20(1), 7–10. doi:10.1057/ejis.2010.55
- Papazoglou, M. P. (2012). Cloud Blueprints for Integrating and Managing Cloud Federations. In M. Heisel (Ed.), *Festschrift: Vol. 7365. Software service and application engineering. Essays dedicated to Bernd Krärmer on the occasion of his 65th birthday* (Vol. 7365, pp. 102–119). Berlin: Springer. doi:10.1007/978-3-642-30835-2_8
- Papazoglou, M. P., & van Heuvel, W.-J. D. (2006). Service-oriented design and development methodology. *International Journal of Web Engineering and Technology*, 2(4), 412. doi:10.1504/IJWET.2006.010423
- Rai, A., Pavlou, P. A., Im, G., & Du, S. (2012). Interfirm IT Capability Profiles and Communications for Cocreating Relational Value: Evidence from the Logistics Industry. *MIS Quarterly*, 36(1), 233–262. Retrieved from <http://dl.acm.org/citation.cfm?id=2208955.2208970>

- Ritchey, T. (2013). General Morphological Analysis - A general method for non-quantified modelling. Retrieved from <http://www.swemorph.com/pdf/gma.pdf>
- Sindhgatta, R., Sengupta, B., & Ponnalagu, K. (2009). Measuring the Quality of Service Oriented Design. In D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, . . . J. Suzuki (Eds.), *Lecture Notes in Computer Science. Service-Oriented Computing* (Vol. 5900, pp. 485–499). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-10383-4_36
- Solakivi, T., Töyli, J., & Ojala, L. (2013). Logistics outsourcing, its motives and the level of logistics costs in manufacturing and trading companies operating in Finland. *Production Planning & Control*, 24(4-5), 388–398. doi:10.1080/09537287.2011.648490
- Steghuis, C. (2006). *Service granularity in SOA-projects: A trade-off analysis*. Retrieved from <http://essay.utwente.nl/57339/>
- Regulation on the harmonisation of certain social legislation relating to road transport. 1, THE EUROPEAN PARLIAMENT 2006.
- United Nations. (2012). *ADR - European Agreement Concerning the International Carriage of Dangerous Goods by Road: Applicable as from 1 January 2013*. New York [etc.]: United Nations.
- Vaquero, L. M., Rodero-Merino, L., Caceres, J., & Lindner, M. (2008). A break in the clouds. *ACM SIGCOMM Computer Communication Review*, 39(1), 50. doi:10.1145/1496091.1496100
- 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(9), 205–224.
- Wang, P., & Li, Z. (2014). Optimization Design Strategies and Methods of Service Granularity. In J. Xu, V. A. Cruz-Machado, B. Lev, & S. Nickel (Eds.), *Advances in Intelligent Systems and Computing. Proceedings of the Eighth International Conference on Management Science and Engineering Management* (Vol. 281, pp. 1445–1455). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-55122-2_124
- Weißenberg, N., & Springer, U. (2014). Cloud Process Modeling for the Logistics Mall-Object-Aware BPM for Domain Experts. *Open Journal of Mobile Computing and Cloud Computing*, 1(2), 31–49.
- Wilding, R., & Juriado, R. (2004). Customer perceptions on logistics outsourcing in the European consumer goods industry. *International Journal of Physical Distribution & Logistics Management*, 34(8), 628–644. doi:10.1108/09600030410557767
- Yang, Y., Pan, S., & Ballot, E. (2016). Mitigating supply chain disruptions through interconnected logistics services in the Physical Internet. *International Journal of Production Research*, 1–14. doi:10.1080/00207543.2016.1223379

Acknowledgement

The work presented in this paper was funded by the German Federal Ministry of Education and Research within the project Logistik Service Engineering und Management (LSEM). More information can be found under the reference BMBF 03IPT504X and on the website www.lsem.de.

Contact Information

Dipl.-Ing. Michael Glöckner
Leipzig University
Endowed Chair of Logistics Information Systems
Grimmaische Straße 12

D-04109 Leipzig

Tel.: +49 (0)341 97 33 704

E-Mail: gloeckner@wifa.uni-leipzig.de

WWW <https://www.wifa.uni-leipzig.de/de/iwi/islog.html>

Privacy-Erhaltung in kollaborativen Geschäftsprozessen in der Cloud

Björn Schwarzbach¹

Institut für Wirtschaftsinformatik, Universität Leipzig¹

Zusammenfassung

Die heutige Wirtschaft organisiert sich zunehmend in Form kollaborativer Netzwerke. Obwohl die Entscheidungsträger die Vorteile der Kollaboration kennen, fürchten sie zugleich einen Verlust der Kontrolle über ihre Daten. Im Rahmen dieses Beitrags wird zunächst die Architektur einer Plattform zur datenschutzerhaltenden Modellierung und Ausführung von kollaborativen Geschäftsprozessen im Cloud-Umfeld vorgestellt. Daran anschließend wird ein Modell und eine Methode zur Modellierung von Datenschutzerfordernungen in kollaborativen Geschäftsprozessen erarbeitet und durch ein Beispiel erläutert.

1 Einleitung

2008 erläuterte der Vizepräsident der Gartner Research Thomas J. Bittman seine Gedanken zur Zukunft des Cloud Computing (CC). So wurden in einer ersten monolithischen Phase Cloud Services durch wenige Anbieter in Form von proprietären Lösungen angeboten. Typische Beispiele hierfür sind Google oder Amazon. In der zweiten Phase hingegen wurden neuartige Dienste entwickelt, welche vertikal entlang der Supply Chain ausgerichtet sind und die proprietären Dienste der ersten Phase verwenden. In der letzten Phase schlossen und schließen sich kleine Cloud-Anbieter zusammen, um gemeinsam komplexere Dienste in Form einer Föderation anbieten zu können. Gleichzeitig setzte sich eine fortschreitende Standardisierung im Bereich CC durch, welche diese Kombination verschiedener Cloud-Dienste zu neuen Diensten begünstigte. (Bittman, 2008) Heutige Cloud-Dienste entsprechen zunehmend dieser dritten Phase. Ein weiterer logischer Schritt hin zu einer vierten Phase ist es, die Dienste mehrerer Clouds, welche nicht zwingend direkte Kopplung unterstützen, zu Ketten zu verbinden.

In vielen Unternehmen werden die internen Abläufe der Supply Chain in Form von Geschäftsprozessen modelliert und teilweise auch automatisiert ausgeführt. Gerade entlang komplexer Supply Chains sind oft viele Unternehmen beteiligt, welche in Form einer Kollaboration oder Kooperation zusammenarbeiten. Diese Unternehmen modellieren gemeinsam einen kollaborativen Geschäftsprozess, in dem sie ihre eigenen Aktivitäten jeweils unter lokaler Kontrolle verantworten. Zur Unterstützung dieser Prozesse liegt die Anwendung des Cloud-Ansatzes nahe. Insbesondere wird durch die Mehrmandantenfähigkeit und die virtuell unerschöpflichen Ressourcen und der hohen Skalierbarkeit der Cloud die Integration mehrerer und auch schnell wechselnder Unternehmen in einem Geschäftsprozess unterstützt.

Während der Ausführung eines Geschäftsprozesses müssen zu im Modell genau definierten Zeitpunkten Informationen zwischen den beteiligten Akteuren ausgetauscht werden. Jede dieser Informationsübertragungen unterliegt dabei speziellen und akteursindividuellen Datenschutzerfordernungen. Solche Datenschutzerfordernungen werden bislang lediglich in Verträgen definiert und nur im Falle einer Datenschutzverletzung, sofern diese auffällt, näher geprüft. In einem solchen Fall sind die finanziellen und rechtlichen Konsequenzen teils beträchtlich, so dass es wünschenswert ist, eine mögliche Datenschutzverletzung zu erkennen und zu unterbinden, bevor sie auftritt.

Das Promotionsprojekt widmet sich daher unter Hilfe der Methoden der gestaltungsorientierten Wirtschaftsinformatik der Frage, wie solche Datenschutzerfordernungen mehrerer Unternehmen in kollaborativen Geschäftsprozessen unter lokaler Kontrolle der beteiligten Unternehmen modelliert werden können, so dass deren Einhaltung sowohl zur Design- als auch zur Runtime überprüft und gesichert wird. Die wesentlichen Artefakte des Projekts sind die Architektur der Plattform zur Modellierung und Ausführung kollaborativer Geschäftsprozesse, das Modell zur Definition der Datenschutzerfordernungen, sowie die Transformation dieses Modells in ein maschinenauswertbares Modell. Die Artefakte betten sich für die Evaluation in einen Gesamtprototypen ein, welcher abschließend die Anwendbarkeit des entwickelten Ansatzes anhand von Beispielprozessen der Logistik in Laborexperimenten demonstriert.

Der Rest dieses Beitrags gliedert sich wie folgt: Im nächsten Kapitel wird die Architektur für die Modellierung und Ausführung kollaborativer Geschäftsprozesse vorgestellt. Im Anschluss daran wird das Modell für die Datenschutzerfordernungen vorgestellt und an einem Beispiel die Transformation in ein maschinenverständliches Modell durchgeführt. Den Abschluss des Beitrags bildet eine Zusammenfassung sowie ein kurzer Ausblick auf die nächsten Schritte im Promotionsprojekt.

2 Architektur zur Modellierung und Ausführung kollaborativer Geschäftsprozesse sowie der beteiligten Datenschutzerfordernissen

Die Grundlage für die Modellierung und die Ausführung von kollaborativen Geschäftsprozessen in der Cloud ist die Plattform, welche die dafür benötigten Dienste bereitstellt. Die Architektur dieser Plattform, welche in *Abbildung 1* aufgezeigt wird, untergliedert sich in drei wesentliche Bereiche: zentrale Plattform, Gateways und Drittanbieterdienste.

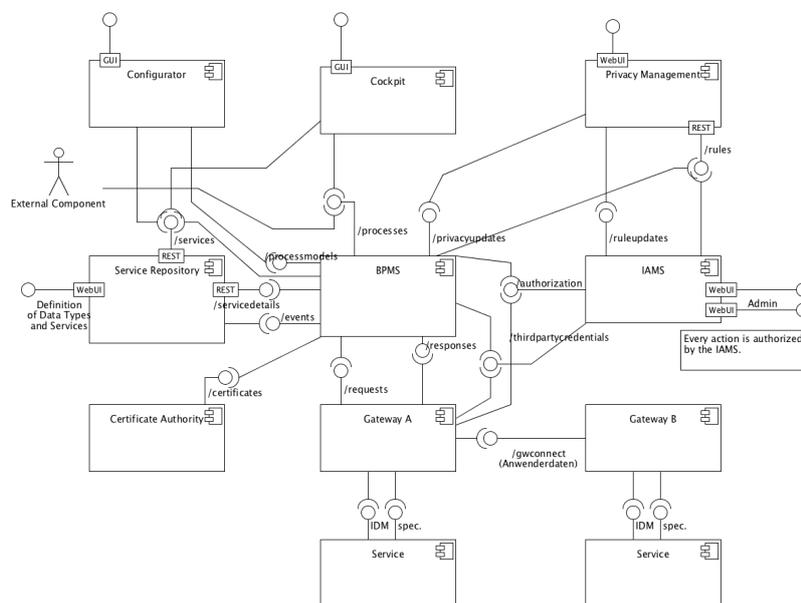


Abbildung 1 Architektur der Plattform zur Modellierung und Ausführung kollaborativer Geschäftsprozesse

Die Aktivitäten eines kollaborativen Geschäftsprozesses können dem Gedanken der Service Orientierung folgend als Businessservice verstanden werden. (Kluge, 2012) entwickelte einen Ansatz, durch den für einen gegebenen Businessservice mögliche elektronische Dienste ausgewählt werden können. Diesem Ansatz folgend wird ein kollaborativer Geschäftsprozess als eine Verknüpfung von elektronischen Diensten verstanden. Diese Dienste können einerseits als on-premis Lösungen bei Unternehmen, andererseits auch als Cloud-Dienste implementiert sein und bilden die unterste Schicht der Architektur.

Diese Dienste werden in einem Konfigurator durch die Anwender, den sogenannten Prozessdesigner, zu Geschäftsprozessen modelliert. Dabei greift der Konfigurator auf die im Service Repository gespeicherten Informationen über die Dienste zurück. In diesem sind alle

Dienste, welche über die Plattform angebunden werden, aufgeführt. Möchte ein Unternehmen einen weiteren Dienst hinzufügen, so kann dieser über eine Web-Schnittstelle eingereicht werden. Nach erfolgter Prüfung des Plattformbetreibers wird der Service zur Nutzung freigegeben. Diese Prüfung umfasst vor allem die Verifikation der gemachten Angaben wie Ausführungsort und Zertifizierungen. Die Dienste sind im Service Repository in Diensttypen unterteilt, wobei Dienstypen jeweils den gleichen Zweck erfüllen und semantisch gleiche Ein- und Ausgabedaten haben. Zu diesem Abgleichzweck wird ein plattforminternes Datenschema verwendet.

Nach der erfolgten Modellierung eines Geschäftsprozesses fügt der Prozessdesigner seine Datenschutzerfordernungen zu dem Prozess sowie den einzelnen Aktivitäten hinzu. Ein Prozess, welcher modelliert und mit Datenschutzerfordernungen versehen ist, wird im Anschluss in das Business Process Management System (BPMS) übertragen und dort gesichert. Sobald ein Geschäftsprozess ausgeführt werden soll, wird von dem Geschäftsprozessmodell eine Instanz erzeugt und dieses zur Ausführung gebracht. Das BPMS prüft bezieht die Entscheidungen über die Erfüllbarkeit von Datenschutzerfordernungen mit in die Service Selection ein und wählt für jede Aktivität den passenden elektronischen Dienst aus.

Das Privacy Management System dient der Definition von Datenschutzerfordernungen von Diensten sowie der Zusammenführung der verschiedenen Datenschutzerfordernungen und der Übertragung dieser an das Identity and Access Management System (IAMS).

Das IAMS entscheidet anhand der übermittelten Datenschutzerfordernungen über deren Erfüllbarkeit für einen speziellen Datenzugriff durch einen elektronischen Dienst im Geschäftsprozess.

Zur Visualisierung des Ablaufs des Geschäftsprozesses sowie der aktuellen globalen und detaillierten Datenschutzsituation wird in einem Cockpit durch verschiedene Ansichten den beteiligten Unternehmen jeweils ein Report erstellt, welcher jederzeit aktuell und umfassend ist.

Die Kommunikation der zentralen Komponenten mit den Drittanbieterdiensten findet ausschließlich durch Gateways statt, welche die mittlere Schicht der Architektur bilden. Gateways werden, durch den Gatewaycontroller erzeugt, sobald das BPMS einen Dienst aufrufen möchte. Jedes Gateway besteht aus vier Komponenten: Communicator, Privacy Guard, Data Adapter und Service Adapter. Während der Communicator für die Sicherheit der Übertragung der Informationen Sorge trägt, indem dieser sämtliche Verbindungen von und zur zentralen Plattform autorisiert, kontrolliert und verschlüsselt, dient der Privacy Guard der Anwendung von Datenschutzauflagen bezüglich der übertragenen Daten. So kann der Privacy Guard beispielsweise Anonymisierung und Pseudonymisierung entsprechend der Datenschutzerfordernungen durchführen. Im Anschluss an den Privacy Guard durchlaufen die Daten den Data Adapter, welcher das Datenformat des Serviceaufrufs vom internen Datenschema in das des Dienstes transformiert, bevor der Service Adapter den eigentlichen Service Call durchführt und somit den Dienst bei dem Dienstanbieter aufruft. Die Antwort des Dienstes durchläuft diese Komponenten des Gateways in umgekehrter Reihenfolge und wird dabei wieder zurück in das interne Datenschema transformiert und gegebenenfalls

depseudonymisiert. Nachdem der Communicator die Antwort auf den Dienstaufwurf an den Gatewaycontroller zurückgegeben hat, zerstört der Gatewaycontroller das gesamte Gateway. Auf diese Weise wird sichergestellt, dass keine Daten zwischen zwei Serviceaufrufen unbeabsichtigt übertragen werden.

Die Kommunikation innerhalb der Plattform einschließlich der Gateways wird durch REST-basierte Webservices realisiert. Jede Komponente der Gateways sowie jede Komponente der zentralen Plattform setzt bei der Autorisation von Verbindungsanfragen auf Client-Zertifikate.

3 Modellierung von Datenschutzerfordernungen

Nachdem die grundsätzliche Architektur der Plattform vorgestellt wurde, wird in diesem Kapitel der entwickelte Ansatz zur Modellierung von Datenschutzerfordernungen für kollaborative Geschäftsprozesse vorgestellt.

Datenschutzerfordernungen entscheiden über die Berechtigung zum Zugriff auf Daten. Daher wurde für die Modellierung der Datenschutzerfordernungen auf Ansätze der Zugriffskontrolle zurückgegriffen.

3.1 Modelle der Zugriffskontrolle

Grundsätzlich werden die bestehenden Modelle der Zugriffskontrolle in vier Typen unterschieden: Mandatory Access Control, Discretionary Access Control, Role Based Access Control und Attribute Based Access Control. Anfang der 70er Jahre des letzten Jahrhunderts wurden erste Modelle zur Zugriffskontrolle entwickelt. Aufgrund der geringen Anzahl von Systemen und Nutzern wurden dabei alle Berechtigungen pro Aktivität und Nutzer festgelegt. Dieser Ansatz wird als Discretionary Access Control bezeichnet. Zusätzlich wurde der Zugriff auf genereller Ebene, z.B. „Zugriff nur vom lokalen Terminal gestattet“, definiert. Diese Mandatory Access Control wird auch in modernen Systemen eingesetzt. So wird beispielsweise unter Linux mit SELinux Mandatory Access Control umgesetzt. (Jin et al., 2012, 2012; Sandhu, Coynek, Feinsteink, & Youmank, 1996)

Ende der 80er und Anfang der 90er Jahre des letzten Jahrhunderts hielten Computer in immer mehr Bereichen Einzug, was zu einer gestiegenen Anzahl an Nutzern führte. Durch diese gestiegene Nutzerzahl war es nicht mehr zweckmäßig, jedem Nutzer individuelle Berechtigungen für alle Daten zuzuweisen. Bei der Role Based Access Control werden daher Nutzer zu Rollen zusammengefasst. Diesen Rollen werden wiederum Berechtigungen zugewiesen. Dabei kann ein Nutzer mehreren Rollen zugewiesen sein. Zusätzlich können neben Nutzern auch Rollen in Rollen organisiert werden, so dass eine Hierarchie mit einer entsprechenden Vererbung von Berechtigungen modelliert wird. Dies vereinfacht die Definition und Wartung von Berechtigungen enorm. Dieser Ansatz wird in den allen aktuellen Windows- und Linux-Systemen sowie vielfältiger weiterer moderner Software eingesetzt. (Ferraiolo & Kuhn, 2009; Ferraiolo, Cugini, & Kuhn, 1995; Iqbal & Noll, 2012; Jin et al., 2012)

Eine der wichtigsten Voraussetzungen für die Eignung eines Ansatzes ist der Einbezug der Zielgruppe in den Erstellungsprozess, z.B. in Form von Experteninterviews während der Analysephase oder durch Experimente zur Evaluation der Erstellten Artefakte. Im Zuge der Auswertung von Experteninterviews aus der Domäne Logistik wurde festgestellt, dass das Verständnis von Datenschutz stark heterogen ist. Zudem wird Datenschutz entgegen der Definition z.B. im Bundesdatenschutzgesetz nicht nur auf personenbezogene oder –beziehbare Daten bezogen, sondern darüber hinaus auch auf geschäftliche Daten. Darüber hinaus zeigte sich in Workshops mit Experten schnell, dass die Anwender keine einheitliche Sicht auf die Definition von Datenschutzerfordernungen haben.

Aufgrund der weiten Verbreitung der Role Based Access Control wurde den Experten zunächst dieser Ansatz vorgestellt. Hierbei stellte sich heraus, dass die Anwender Nutzer in Gruppen einteilen, jedoch diese Gruppen nicht nach strengen Regeln, sondern je nach Anwendungsfall unterschiedlich und teils auch auf Basis von weiteren Eigenschaften der Nutzer gebildet werden. So kann beispielsweise eine Anforderung als „Der Fahrer ist berechtigt, die Telefonnummer des Empfängers des Pakets zu lesen, um diesen anzurufen, sofern er sich gerade in der Stadt des Empfängers aufhält.“ definiert werden. Diese Anforderung enthält die folgenden Informationen:

Der Nutzer, welcher Zugriff auf die Telefonnummer des Empfängers erhalten möchte,

- ist der Rolle Fahrer zugewiesen,
- muss sich in der Stadt des Empfängers befinden und
- der Zweck des Datenzugriffs ist ein Anruf des Empfängers.

Solch eine Anforderung führt zu einer stark steigenden Komplexität der Rollenstruktur der Role Based Access Control, weshalb die Role Based Access Control nicht geeignet für die Modellierung von Datenschutzerfordernungen in kollaborativen Geschäftsprozessen erscheint. Dies wird umso deutlicher, desto mehr Unternehmen an einem solchen Geschäftsprozess beteiligt sind und unterschiedliche Anforderungen beitragen.

Um diese Anforderungen abbilden zu können, wurden im Wesentlichen zwei Ansätze verfolgt. (Gouglidis & Mavridis, 2012) verfolgen die Role Based Access Control, indem das Modell um Aspekte für virtuelle Organisationen erweitert wurde. Das Modell umfasst jedoch keine Geschäftsprozesse, Workflows oder CC. Weitere Ansätze beziehen Geschäftsprozesse ein, lassen jedoch Kooperationen und Kollaborationen außen vor. (Le & Wang, 2012; Le Hung, Doll, Barbosu, Luque, & Wang, 2012, 2014) haben ein Modell vorgestellt und evaluiert, welches ebenfalls die Role Based Access Control erweitert und dabei Team-Kollaboration und Workflows im Gesundheitssektor mit einbezieht. Jedoch bietet auch dieses Modell nicht die erwähnte benötigte Flexibilität zur Definition von Anforderungen.

Um der gestiegenen Komplexität der Anforderungen an die Zugriffskontrolle gerecht zu werden, wurde die Attribute Based Access Control eingeführt, welche die Definition von Regeln zur Zugriffskontrolle auf Basis von Attributen von Subjekten und Objekten ermöglicht. (Hu et al., 2013) definieren die Attribute Based Access Control wie folgt:

„An access control method where subject requests to perform operations on objects are granted or denied based on assigned attributes of the subject, assigned attributes of the object,

environment conditions, and a set of policies that are specified in terms of those attributes and conditions.“ (Hu et al., 2013)

Die Attribute Based Access Control bezeichnet die eine Berechtigung anfragende Entität (z.B. Nutzer, Programm) als Subjekt. Typische Attribute von Subjekten sind deren Name oder deren Position in der Organisation. Die Daten, auf welche Subjekte zugreifen möchten, werden als Ressourcen bezeichnet. Eine Regel zur Zugriffssteuerung wird als Tupel auf Subjekt, Ressource, Operation (z.B. lesen, schreiben), Ergebnis (z.B. erlaubt, verweigert) und eventuell Bedingungen dargestellt. Die bereits vorgestellte Anforderung „Der Fahrer ist berechtigt, die Telefonnummer des Empfängers des Pakets zu lesen, um diesen anzurufen, sofern er sich gerade in der Stadt des Empfängers aufhält.“ kann mit Hilfe der Attribute Based Access Control wie folgt zerlegt werden:

- Subjekt: Der Fahrer
- Ressource: Die Telefonnummer des Empfängers
- Operation: Lesen
- Bedingungen: Aktuelle Aktivität ist „rufe den Empfänger an“; Aufenthaltsort des Fahrers liegt in der Stadt des Empfängers
- Ergebnis: Gestattet

(Jin et al., 2012) zeigt, dass alle mit früheren Modellen definierten Zugriffskontrollregeln in Attribute Based Access Control-Regeln transformiert werden können. Aufgrund der sehr hohen Flexibilität wurde die Attribute Based Access Control als Grundlage für das Modell zur Definition von Datenschutzerfordernungen in kollaborativen Geschäftsprozessen gewählt.

3.2 Modellierung der Datenschutzerfordernungen

Nachdem im letzten Abschnitt die Wahl des zugrundeliegenden Modells erörtert wurde, wird folgend die Modellierung der Datenschutzerfordernungen im Kontext kollaborativer Geschäftsprozesse beschrieben.

In der beschriebenen Plattform werden die Nutzer zwei grundlegenden Rollen zugeordnet. Der Prozessdesigner modelliert mittels den Geschäftsprozess und hat daher das Verständnis für den Gesamtprozess. Der Dienstanbieter hingegen stellt in der Plattform seine elektronischen Dienste zur Nutzung in den Geschäftsprozessen bereit, hat jedoch keine Kenntnis über die Geschäftsprozesse.

Beide Rollen haben den Anspruch, Datenschutzerfordernungen definieren zu können. Dabei wird davon ausgegangen, dass Datenschutzerfordernungen stets die Ausgabedaten eines elektronischen Dienstes betreffen. Die eingangs erwähnte Anforderung wird an dem Dienst definiert, der die Telefonnummer des Empfängers bereitstellt. Neben den Datenschutzerfordernungen beider Rollen, Prozessdesigner und Dienstanbieter, werden zusätzlich Gesetze und Richtlinien durch den Plattformbetreiber bereitgestellt. Die drei Bereiche sind in *Abbildung 2* veranschaulicht.

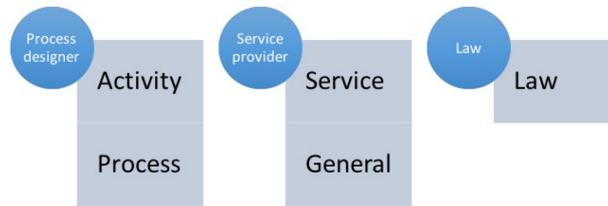


Abbildung 2 Sichten der Datenschutzanforderungen

Um den Prozess der Definition der Datenschutzanforderungen für die Akteure zu vereinfachen und somit einer Ermüdung vorzubeugen, werden die Anforderungen für beide Rollen jeweils in eine generelle und eine spezielle Sicht unterteilt.

Der Prozessdesigner definiert auf Prozessebene Anforderungen, welche für alle Aktivitäten eines Prozesses Gültigkeit haben. Diese generellen Anforderungen können für einzelne Aktivitäten des Prozesses durch spezifische Regeln übersteuert werden. Dabei hat jeweils eine aktivitätsbezogene Regel vor einer prozessbezogenen Regel Vorrang.

Der Dienstanbieter definiert zunächst allgemeine Datenschutzanforderungen, welche für alle von ihm angebotenen Dienste zutreffen. Für einzelne Dienste kann der Dienstanbieter diese Regeln durch dienstspezifische Regeln überschreiben. Auch bei dieser Rolle haben die dienstbezogenen Regeln Vorrang vor den generellen Regeln.

In den durchgeführten Workshops zeigte sich, dass die Logistiker aktuell noch gruppenbasiert denken, jedoch zunehmend weitere Bedingungen wie Zeitabhängigkeit hinzufügen. Die Gruppen sollen dabei von jedem Nutzer unabhängig definiert werden können. Um dies zu ermöglichen, wurde in einem ersten Prototyp eine Darstellung in Tabellenform gewählt, wobei die Spaltenköpfe die vom Nutzer definierten Gruppen (Subjekte) und die Zeilen die Ressourcen spezifizieren. In den Zellen wird vom Anwender jeweils angegeben, ob der Zugriff gewährt oder verweigert werden soll. Den Gruppen kann jeweils ein Ablaufdatum zugewiesen werden, so dass beispielsweise spezielle Berechtigungen für die Dauer eines Vertrags definiert werden können. Eine entsprechende Tabelle ist in dargestellt.

In den Zellen kann neben Erlaubt und Verboten Keine Angabe gewählt werden. Keine Angabe bewirkt, dass die nächst allgemeinere Definition zutrifft.

Attribut	Standard	Foxi ▶07.10.2016
▼ Services	⊘ Verboten ▼	✓ Erlaubt ▼
▼ WarehouseCapacities	⊘ Verboten ▼	✓ Erlaubt ▼
FreeWarehouseId	⊘ Verboten ▼	✓ Erlaubt ▼
CompanyName	⊘ Verboten ▼	✓ Erlaubt ▼
ServiceName	⊘ Verboten ▼	✓ Erlaubt ▼
FreeBed	⊘ Verboten ▼	✓ Erlaubt ▼
FreeWeight	⊘ Verboten ▼	✓ Erlaubt ▼
FreeVolume	⊘ Verboten ▼	✓ Erlaubt ▼
FromDate	⊘ Verboten ▼	✓ Erlaubt ▼
ToDate	⊘ Verboten ▼	✓ Erlaubt ▼
▶ TransportCapacities	⊘ Verboten ▼	✓ Erlaubt ▼

Abbildung 3 Tabelle mit Datenschutzerfordernissen eines Dienstbieters zweier Dienste mit zwei Gruppen und Ablaufdatum

Bei der Definition der Gruppen wird o.B.d.A. aktuell nur der Firmenname unterstützt. Dies kann jedoch auf einfache Weise um beliebige Attribute ergänzt werden.

3.3 Auswertung der Datenschutzerfordernissen

In diesem Abschnitt wird der Algorithmus zur Überführung der im vorangegangenen Abschnitt beschriebenen Datenschutzerfordernissen in eine maschinenverständliche und ausführbare Form vorgestellt und kurz an einem Beispiel erläutert.

Zur Auswertung von attributbasierten Zugriffsregeln eignet sich insbesondere der XACML-Standard. XACML ist seit 2013 in Version 3.0 verfügbar und findet eine zunehmende Verbreitung, (OASIS, 2013) so dass die Verwendung von XACML als Basis für die Datenschutzerfordernissenauswertung geeignet scheint. Es existiert eine Vielzahl von Werkzeugen für XACML, welche jedoch in den wenigsten Fällen mit der aktuellen Version des Standards kompatibel sind. AT&T hat seine Implementierung des XACML 3.0 Standards als Apache Incubator Projekt bereitgestellt, welches jedoch seitdem eingefroren und nicht lauffähig ist. (AT&T, 2015)

Den Kern der Architektur von XACML 3.0 bildet ein Policy Decision Point (PDP). Dieser wertet eine Anfrage anhand von Regeln, welche in einem Policy Administration Point (PAP) erstellt wurden, aus. Sollten für die Auswertung Attribute benötigt werden, welche in der Anfrage nicht übermittelt wurden, werden diese bei einem passenden Policy Information Point (PIP) angefragt. Auf die beschriebene Plattform für kollaborative Geschäftsprozesse bezogen ist das IAMS der PDP und ein PIP, das PMS ist der PAP. Die anfragende Komponente wird in XACML als Policy Enforcement Point (PEP) bezeichnet, was in der beschriebenen Architektur dem BPMS und dem Privacy Guard entspricht.

XACML ist neben einer Architektur eine Sprache zur Modellierung von Zugriffsregeln. Die Sprache ist xml-basiert und spezifiziert drei Hauptelemente. Das PolicySet enthält Policies und PolicySets, wohingegen eine Policy Rules beinhaltet. Eine Rule ist eine einzelne Aussage mit einem bestimmten Effekt (meist permit oder deny) sowie keiner, einer oder mehrerer Bedingungen. Jedes Element enthält zudem ein Target, welches die Anwendbarkeit des entsprechenden Elements spezifiziert.

Die Transformation der modellierten Datenschutzerfordernungen in XACML erfolgt in vier Schritten:

- Erzeugen von Datenschutzerfordernungen für alle Dienste und Aktivitäten eines Prozesses
- Auflösung aller „Keine Angabe“ Einträge zu „Erlaubt“ oder „Verboten“
- Übersetzung jeder Datenschutzerforderung (je Dienst und Aktivität) in XACML-Fragmente
- Zusammensetzen der XACML-Fragmente zu einem Dokument

Das Diagramm zeigt die Transformation von globalen Datenschutzerfordernungen in spezifische XACML-Gruppen. Die oberste Tabelle (orange) zeigt die globale Sicht mit 'N/S' (Keine Angabe) für 'GoodRelations' und 'NeverAgain'. Die mittlere Tabelle (gelb) zeigt die Auflösung dieser Einträge zu 'Permit' oder 'Deny'. Die unterste Tabelle (gelb) zeigt die resultierende XACML-Struktur mit definierten Gruppen.

Attribute	Default	GoodRelations	NeverAgain
Address			
Street	Deny	Permit	Deny
Zipcode	Permit	N/S	Deny
City	Permit	N/S	N/S

Attribute	Default	GoodRelations	NeverAgain
Address			
Street	Deny	Permit	Deny
Zipcode	N/S	Permit	Deny
City	Permit	Permit	N/S

Attribute	Default	GoodRelations	NeverAgain
Address			
Street	Deny	Permit	Deny
Zipcode	Permit	Permit	Deny
City	Permit	Permit	Permit

Abbildung 4 Auflösung von „Keine Angabe“ Einträgen der Datenschutzerfordernungen (Permit: Erlaubt, Deny: Verboten, N/S: Keine Angabe)

Ein Beispiel des der Durchführung des dritten Schrittes ist in dargestellt. Es wurden neben der Default-Gruppe zwei weitere Gruppen GoodRelations und NeverAgain definiert. Die oberste Tabelle stellt die globale Sicht der Datenschutzerfordernungen dar, welche für einen speziellen Service durch die mittlere Tabelle überschrieben werden. Beispielhaft wird der Zugriff auf die

Ressource City durch eine Firma der Gruppe NeverAgain mit keine Angabe in der dienstspezifischen Tabelle spezifiziert. Dies führt zur Anwendung des Eintrags der Default Gruppe und resultiert somit in Permit („Erlaubt“).

Die Spaltenköpfe der Tabelle werden in XACML direkt als Subjekte modelliert, die Attribute werden als Ressourcen umgesetzt. Für jede Zelle wird eine Rule erzeugt. Die Tabelle wird schließlich als PolicySet modelliert, welches ein Policy je Attribut und diese wiederum je eine Policy je Gruppe enthält.

Im vierten Schritt werden die erstellten PolicySet in einem PolicySet zusammengefasst.

4 Zusammenfassung und Ausblick

In diesem Beitrag wurde eine Architektur zur Modellierung und Ausführung von kollaborativen Geschäftsprozessen dargestellt, welche während der Modellierung und Ausführung der Prozesse den Datenschutz sicherstellt. Im Anschluss wurde ein für die Definition der dafür benötigten Datenschutzanforderungen geeignetes Modell vorgestellt und die Methode zur Umsetzung dieser Datenschutzanforderungen erörtert.

Die vorgestellten Artefakte befinden sich momentan überwiegend in der Evaluation und haben dabei überwiegend ihre Eignung erwiesen. Insbesondere die Definition der Datenschutzanforderungen in mehreren Sichten wurde als zielführend erachtet. Hingegen eignet sich die Tabellendarstellung der Regeln nur begrenzt für umfangreiche und komplexe Anforderungen. Hierfür wird in der Arbeitsgruppe an weiteren Darstellungsvarianten gearbeitet. Da der Fokus des Promotionsprojekts jedoch auf der Architektur der Plattform und der Auswertung der modellierten Regeln liegt, wird dies Gegenstand weiterer Arbeiten sein.

References

- AT&T. (2015). *At&t xacml 3.0 implementation*. Retrieved from <https://github.com/att/XACML>
- Bittman, T. (2008). *The Evolution of the Cloud Computing Market*. Retrieved from http://blogs.gartner.com/thomas_bittman/2008/11/03/the-evolution-of-the-cloud-computing-market/
- Ferraiolo, D. F., & Kuhn, D. R. (2009). Role-based access controls. *arXiv preprint arXiv:0903.2171*,
- Ferraiolo, D., Cugini, A., & Kuhn, R. (1995). Role Based Access Control: Features and Motivation. In *Computer Security Application Conference* .
- Gouglidis, A., & Mavridis, I. (2012). domRBAC: An access control model for modern collaborative systems. *computers & security*, 31(4), 540–556.
- Hu, V., Ferraiolo, D., Kuhn, R., Schnitzer, A., Sandlin, K., Miller, R., & Scarfone, K. (2013). *Guide to Attribute Based Access Control Definition and Considerations*: NIST.
- Iqbal, Z., & Noll, J. (2012). Towards semantic-enhanced attribute-based access control for cloud services. In *2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications* (pp. 1223–1230).

- Jin, X., Krishnan, R., & Sandhu, R. (2012). A unified attribute-based access control model covering DAC, MAC and RBAC. In *IFIP Annual Conference on Data and Applications Security and Privacy* (pp. 41–55).
- Kluge, R. (2012). *Preselection of Electronic Services by Given Business Service Based on Measuring Semantic Heterogeneity within the Application Area of Logistics*. Leipzig, Univ., Diss., 2012. Retrieved from <http://nbn-resolving.de/urn:nbn:de:bsz:15-qucosa-98543>
- Le, X. H., & Wang, D. (2012). Development of a system framework for implementation of an enhanced role-based access control model to support collaborative processes. In *Proc 3rd USENIX Workshops on Health Security and Privacy*.
- Le Hung, X., Doll, T., Barbosu, M., Luque, A., & Wang, D. (2012). An enhancement of the Role-Based Access Control model to facilitate information access management in context of team collaboration and workflow. *Journal of Biomedical Informatics*, 45(6), 1084–1107. doi:10.1016/j.jbi.2012.06.001
- Le Hung, X., Doll, T., Barbosu, M., Luque, A., & Wang, D. (2014). Evaluation of an Enhanced Role-Based Access Control model to manage information access in collaborative processes for a statewide clinical education program. *Journal of Biomedical Informatics*, 50, 184–195. doi:10.1016/j.jbi.2013.11.007
- OASIS. (2013). *eXtensible Access Control Markup Language (XACML) Version 3.0: OASIS Standard*. Retrieved from <http://docs.oasis-open.org/xacml/3.0/xacml-3.0-core-spec-os-en.html>
- Sandhu, R. S., Coynek, E. J., Feinstein, H. L., & Youmank, C. E. (1996). Role-based access control models yz. *IEEE computer*, 29(2), 38–47.

Kontaktinformationen

Dipl.-Inf. Björn Schwarzbach
Universität Leipzig
Fakultätsrechenzentrum
Grimmaische Straße 12
D-04109 Leipzig

Tel.: +49 (0)341 97-33615

E-Mail: schwarzbach@wifa.uni-leipzig.de

Gewichtung von Anforderungen an ein Process Performance Measurement System zur Unterstützung einer ganzheitlichen Working-Capital-Steuerung

Marcus Pfitzner, Prof. Dr. Andreas Hilbert

Fakultät für Wirtschaftswissenschaften, Lehrstuhl für Wirtschaftsinformatik | Business Intelligence Research, Technische Universität Dresden

Zusammenfassung

Ziel dieses Beitrages ist die Gewichtung von Anforderungen an ein Process Performance Measurement System zur Unterstützung einer ganzheitlichen Working-Capital-Steuerung. Die zuvor über ein Literature Review erhobenen Anforderungen werden beschrieben und mittels der Methode des Analytical Hierarchy Process gewichtet. Als Ergebnis wird eine kundenorientierte Priorisierung der Anforderungen an ein PPMS in diesem Kontext zur Verfügung gestellt. Diese Gewichtung legt die Grundlage für spätere Forschungsarbeiten hinsichtlich einer theoretischen Konzeption und prototypischen Implementierung eines derartigen PPMS.

1 Einleitung

Die effiziente Nutzung und Stärkung der eigenen Innenfinanzierungskraft stellt für Unternehmen einen erheblichen Faktor zur Wettbewerbsfähigkeit dar (Klepzig & Vater, 2015). Spätestens seit der letzten Finanzkrise und den einhergehenden restriktiveren Kreditvergaben sowie den volatileren Finanzmärkten rücken Liquiditätssicherung und Ertragsstabilisierung in den Fokus (Hofmann & Martin, 2014; Lind, Pirttilä, Viskari, Schupp, & Kärr, 2012). Die Liquiditätsfreisetzung aus dem operativen Geschäft heraus, die bei

gewinnmaximierender Verwendung der freigesetzten Mittel auch einen Beitrag zur Rentabilitätssteigerung leistet (Aktas, Croci, & Petmezas, 2015; Knauer & Wöhrmann, 2013), kann aus einer gezielten Steuerung des Working Capital (WoC) hervorgehen (Sure, 2015). Allerdings werden die liquiditäts- und rentabilitätssteigernden Effekte eines nachhaltigen Working Capital Management (WCM) nur unzureichend genutzt, da bisher kein ganzheitlicher Ansatz zur Steuerung des WoC existiert (Döring, Schönherr, & Steinhäuser, 2012; Pfitzner & Hilbert, 2014). Die Gründe dafür sind unter anderem in der hohen Komplexität des zugrundeliegenden Prozessgefüges, der Menge an Daten und der Vielfältigkeit an Datenquellen zu sehen, die für eine effiziente WoC-Steuerung zu beachten und einzubeziehen sind (Enqvist, Graham, & Nikkinen, 2014). Um dieser Komplexität der beteiligten Prozesse zu genügen, bedarf es eines geeigneten, IT-gestützten Werkzeugs zur Beurteilung ihrer Leistungsfähigkeit (Hofmann & Martin, 2014; Protopappa-Sieke & Schimke, 2010). In diesem Zusammenhang kann ein Process Performance Measurement System (PPMS) einen möglichen Lösungsansatz bieten (Kueng, 2000; Pfitzner & Hilbert, 2014). Basierend auf der vorangegangenen Forschung der Autoren zu Anforderungen an ein solches System, soll nun die Relevanz der erhobenen Anforderungen bestimmt werden. Ziel dieses Beitrages ist demnach die Anforderungsgewichtung für ein PPMS zur Unterstützung einer ganzheitlichen WoC-Steuerung. Zu diesem Zweck werden die Methode Analytic Hierarchy Process (AHP) angewendet und Expertenworkshops durchgeführt.

In den nachfolgenden Kapiteln werden zunächst die Grundlagen des WCM betrachtet sowie die Anforderungen an ein PPMS zur Unterstützung einer ganzheitlichen WoC-Steuerung beschrieben. Daran anknüpfend erfolgt die Gewichtung der Anforderungen vorgenommen, die im Ergebnis eine kundenorientierte Priorisierung der Anforderungen an ein PPMS zur Unterstützung einer ganzheitlichen WoC-Steuerung ergibt. Die Ergebnisse legen die Grundlage für eine spätere theoretische Konzeption und prototypische Implementierung eines derartigen PPMS. Der Artikel schließt mit einer Zusammenfassung der wesentlichen Ergebnisse und gibt einen Ausblick auf das weitere Forschungsvorhaben.

2 Working Capital Management

Das WCM zielt auf eine Reduzierung des gebundenen Umlaufvermögens zur Liquiditätsgenerierung und Rentabilitätssteigerung ab. Durch Senkung der kurzfristigen Forderungen sowie des Vorratsbestandes und der Verlängerung bestehender Verbindlichkeiten werden finanzielle Mittel freigesetzt, die gewinnbringend investiert werden können. Aus diesen Hauptkomponenten des WoC abgeleitet sind die drei zugrundeliegenden Teilprozesse, „Order-to-Cash“ (Forderungsmanagement), „Forecast-to-Fulfill“ (Vorrätemanagement) und „Purchase-to-Pay“ (Verbindlichkeitsmanagement), zu betrachten (Döring et al., 2012). Das Ziel einer ganzheitlichen Steuerung ist es, das WoC bzw. den Cash-Conversion-Cycle, der die Dauer der Kapitalbindung als Zeitspanne der Umwandlung eines Geldabflusses in einen Geldzufluss misst, unter Berücksichtigung bestehender Wechselwirkungen zu minimieren. Zur Erreichung dieses Ziels müssen zahlreiche Prozesse und Datenströme im Unternehmen berücksichtigt werden (Sure, 2015), um den jeweiligen

Beteiligten adäquate entscheidungsrelevante Informationen zur Verfügung zu stellen. Diese Aufgabe kann nur mit Hilfe eines geeigneten IT-basierten Werkzeugs bewältigt werden. Verschiedene Beiträge zeigen, dass es derzeit noch an einem ganzheitlichen System mangelt (Rupp, 2011). Einen möglichen Ansatz kann in diesem Kontext ein PPMS als ganzheitliches System zur Erhebung und Bewertung relevanter Performancedaten sowie zur Verteilung der erlangten Ergebnisse an die Prozessbeteiligten bieten (Kueng, 2000; Pfitzner & Hilbert, 2014). Basierend auf einer State-of-the-art-Analyse bezüglich der PPMS-Forschung konstatiert Glavan, dass es bisher keine Entwicklung eines integrierten, ganzheitlichen PPMS seit der Einführung des Konzeptes nach Kueng gegeben hat (Glavan, 2011; Kueng, 2000). Darauf aufbauend haben die Autoren in einer vorgelagerten Forschungsarbeit mittels eines systematischen Literature Review sowie unter Anwendung der Methode des Quality Function Deployment (QFD) Anforderungen an ein solches System identifiziert und Konstruktionsmerkmale abgeleitet (Wieland, Pfitzner, Fischer, & Hilbert, 2015). Die Auswahl der zugrunde gelegten wissenschaftlichen Literatur fußt auf einer Schlüsselwort-basierten Suche in wissenschaftlichen Journalen, die von der Wissenschaftlichen Kommission Wirtschaftsinformatik mit „A“ und „B“ bewertet wurden. Dazu wurden die Websites der jeweiligen Journale bzw. wissenschaftliche Literaturdatenbanken genutzt. Die ermittelten Anforderungen sollen nun im Kontext einer ganzheitlichen WoC-Steuerung gewichtet werden, um die Basis einer kundenorientierten Lösung zu schaffen.

3 Analytic Hierarchy Process

Zur Gewichtung der Anforderungen an ein PPMS mit dem Ziel der Unterstützung einer ganzheitlichen WoC-Steuerung wird die Methode AHP verwendet. AHP kommt grundlegend für Anwendungsfälle zum Einsatz, die mit einem multikriteriellen Entscheidungsproblem vergleichbar sind (Mühlbacher & Kaczynski, 2013; Saaty, 2003). Die Methode zielt darauf ab subjektive Entscheidungen zu überprüfen und zu ergänzen und so mit möglichst niedrigem Zeitaufwand tragbare Lösungen zu finden. Ein Augenmerk liegt vor allem auf der Nachvollziehbarkeit der Entscheidungsfindung und des Ergebnisses sowie der Identifizierung eventueller Inkonsistenzen in der Entscheidungsfindung (Mühlbacher & Kaczynski, 2013). Somit ist AHP im vorliegenden Anwendungsfall (Anforderungsgewichtung an ein IT-System) eine geeignete Methode zur Erreichung des Ziels dieses Beitrags.

Im AHP werden Kriterien, die zur Lösung eines Problems herangezogen werden, stets in eine hierarchische Struktur gebracht („Hierarchy“). Elemente einer Hierarchie können dabei in Gruppen eingeteilt werden, wobei jede Gruppe nur jeweils eine andere („höhere“) Gruppe von Hierarchieelementen beeinflusst und nur von einer anderen („niedrigeren“) beeinflusst wird. AHP ist darüber hinaus geeignet, eine Problemkonstellation in all ihren Abhängigkeiten umfassend zu analysieren („Analytic“). Darüber hinaus gibt die Methode einen prozessualen Ablauf vor, wie Entscheidungen strukturiert und analysiert werden („Process“). Dieser Ablauf ist weitgehend gleichbleibend, wodurch der AHP bei mehrfachem Einsatz zu einem leicht einsetzbaren Entscheidungswerkzeug wird. Am Anfang des AHP stehen zunächst die Definition des Entscheidungsproblems sowie die Aufstellung einer Zielhierarchie im

Vordergrund. Im Anschluss daran werden die Prioritäten auf Basis der erhobenen Präferenzurteile in verschiedenen Vergleichsmatrizen zusammengefasst. Darauf aufbauend erfolgen die Matrizenberechnung sowie die Bestimmung der Gewichtungsfaktoren auf Basis der Eigenvektorberechnung. In Verbindung mit der abschließenden Berechnung der Gesamthierarchie ergibt sich eine kundenorientierte Priorisierung der Anforderungen (Meixner & Haas, 2002; Saaty, 2003).

4 Anforderungsgewichtung eines PPMS zur Unterstützung einer ganzheitlichen WoC-Steuerung

In diesem Kapitel werden die Anforderungen an ein PPMS zur Unterstützung einer ganzheitlichen WoC-Steuerung beschrieben, welche in einem nächsten Schritt mittels AHP gewichtet werden.

4.1 Beschreibung der Anforderungen an ein PPMS

Für eine ganzheitliche und kundenorientierte Gestaltungsempfehlung für ein PPMS wurden in der vorherigen Forschung konkrete Anforderungen und Konstruktionsmerkmale eines PPMS identifiziert und beschrieben, die sich als Konstruktionsgrundlage für einen IT-technischen Systementwurf eignen. Im Folgenden finden die Anforderungen, die in diesem Beitrag im Kontext einer ganzheitlichen WoC-Steuerung gewichtet werden sollen, eine kurze Betrachtung.

CR	Bezeichnung der Anforderung	Beschreibung der Anforderung
<i>Messen, Erfassen und Generieren von Daten</i>		
1	Inhalte von KPIs	KPIs müssen quantitativ messbar sein und sind dahingehend ggf. anzupassen. Zur Abbildung einer ganzheitlichen Betrachtungsweise können mehrere KPIs unterschiedlicher Gesichtspunkte, wie bspw. finanzielle oder qualitätsbezogene Aspekte, Kundenzufriedenheit, Kapazitätsauslastung, Anwendung finden (Cleven, 2012). Außerdem bedarf es einer regelmäßigen Überprüfung und ggf. Anpassung der KPIs in Folge eines dynamischen Prozessumfeldes.
2	Identifikation von KPIs	Als essenzieller Aspekt eines PPMS sollten KPIs gewissenhaft und in enger Zusammenarbeit mit den Prozessverantwortlichen und -beteiligten definiert werden (Kueng, 2000; Kueng & Krahn, 1999).
3	Qualität von KPIs	Neben der Eignung zum dauerhaften Monitoring weisen hochwertige KPIs weitere Merkmale auf, wie z.B. Reliabilität, Linearität sowie Sensitivität (Kueng, 2000; Xirogiannis & Glykas, 2004). Des Weiteren müssen die definierten KPIs von den Prozessbeteiligten akzeptiert und für sie nachvollziehbar sein sowie die übergreifende Unternehmensstrategie berücksichtigen (Buchheim, 2000; Kueng, 2000).
4	Umgang mit Dynamik und Veränderungen	Der oftmals im Prozessumfeld zu beobachteten Dynamik geschuldet, ist eine regelmäßige Überprüfung und ggf. Anpassung der KPIs erforderlich (del-Río-Ortega, Resinas, Cabanillas, & Ruiz-Cortés, 2013). In diesem Zusammenhang kommt dem Prozesscontrolling die Aufgabe zu, Anpassungsbedarfe zu identifizieren (Grob, Bensberg, & Coners, 2008). Mit zunehmendem Detaillierungsgrad steigt auch der Anpassungsbedarf der KPIs (Kueng, 2000).

5	weitgehende Nutzung existierender Informationsobjekte	Im Unternehmen ggf. bereits vorhandene Measurement-Systeme und verwendete Daten und KPIs sind in der Entwicklung eines PPMS zu berücksichtigen und hinsichtlich Ihrer Eignung für das angestrebte Ziel zu bewerten (Grosswiele, Röglinger, & Friedl, 2013; Kueng, 2000).
6	Identifikation der relevanten Datenquellen	Damit kein unnötiger Aufwand in der Datenverarbeitung entsteht, sollten nur die Datenquellen Verwendung finden, die für den jeweiligen Anwendungsfall als relevant identifiziert wurden (Cleven, 2012; Kueng, 2000).
<i>Analysieren und Bewerten</i>		
7	Identifikation von Zusammenhängen	Ein PPMS soll die Identifikation bestehender Beziehungen zwischen Kennzahlen und Zielen ermöglichen. Dabei sind mehrere Performance-Aspekte in der Betrachtung zu kombinieren, um empirische Ursache-Wirkungsbeziehungen aufzudecken (Gladen, 2002).
8	Untersuchungsmöglichkeiten auf verschiedenen Detaillierungsstufen	Zur Ursachenanalyse signifikanter Performanceabweichungen soll ein PPMS ein Kennzahlensystem abbilden, das eine Performancebetrachtung auf verschiedenen Detaillierungsstufen erlaubt (Kueng, 2000). Außerdem ist ein logischer Zusammenhang zwischen den unterschiedlichen Level zu gewährleisten.
9	Entscheidungsunterstützung	Den jeweiligen Prozessbeteiligten sollen entscheidungsrelevanter Informationen in optimaler Informationsqualität und -menge bereitzustellen (Wettstein, Kueng, & Meier, 2001). Dazu müssen die jeweiligen Analyseziele und Informationsbedarfe genau beachtet werden (Grosswiele et al., 2013).
10	Kopplung KPI und Prozessschritt	Die Definition von KPIs soll prozessspezifisch erfolgen (Kueng, 2000), da diese Kopplung eine bessere Nachvollziehbarkeit eines Prozesses ermöglicht.
11	Identifikation von Ursachen	Als wesentliches Ziel eines PPMS gestaltet sich die Ursachenanalyse als ein fortlaufender Prozess, der durch die Beseitigung des bedeutendsten Fehlers den Fokus auf den nächsten Fehler legt (Buchheim, 2000).
12	Fehlervermeidung	Es ist eine „valide und konsistente Steuerungslogik zur zielsetzungsgerechten Prozessausführung“ (Grob et al., 2008) zu implementieren, um Fehler im Prozessablauf zu vermeiden. Außerdem ist ein mindestens notwendiges Set an Indikatoren zu bestimmen, um nicht akzeptable Performanceabweichungen zu erkennen (Robson, 2004).
<i>Visualisieren und Verteilen von Informationen / Handlungsbedarf</i>		
13	PPMS muss Prozessverbesserung visualisieren	Definierte KPIs müssen innerhalb eines Measurement-Systems ebenfalls auf Verbesserung fokussiert sein (Kueng, 2000). Dadurch kann deren Entwicklung bezüglich einer Verbesserung in regelmäßigen Berichten kontrolliert werden (Buchheim, 2000). Zur umfassenden Beurteilung möglicher Maßnahmen können die ermittelten Ursache-Wirkungsbeziehungen dienen (Gladen, 2002).
14	Integration von PPMS in den Management-Controlling-Prozess	Damit der Management-Controlling-Prozess ausreichend unterstützt werden kann, ist die Akzeptanz der definierten KPIs auf Managementebene (Kueng, 2000) sowie deren Integration in Controlling-Prozess notwendig (Robson, 2004). Dies vorausgesetzt, ist durch regelmäßige Berichte geeigneter KPIs die Definition neuer Ziele und ggf. eine Strategieanpassung möglich (Kudernatsch, 2002).
15	PPMS muss für interne und externe Zielgruppen nutzbar sein	Definierte KPIs müssen für interne und externe Adressaten verständlich sein, sodass sie Akzeptanz finden (del-Rio-Ortega et al., 2013). Daher müssen KPIs diesbezüglich optimiert und entsprechend der Zielgruppe auf das Passende und Wesentliche reduziert werden (Xirogiannis & Glykas, 2004).
16	einfache, transparente Darstellung der Situation	Die aktuelle Prozessperformance sowie die identifizierten Ursache-Wirkungsbeziehungen sind transparent und leicht verständlich darzustellen (Buchheim, 2000).
17	PPMS muss Prozesssteuerung unterstützen	Die Auswertung relevanter Performancegrößen zur Prozesslaufzeit ist Voraussetzung für eine wirksame Unterstützung der Prozesssteuerung. Die Schnittstellen betroffener Systeme sind daher auf die aktuell gültigen Prognosemodelle und die Ermittlung aktueller Werte anzupassen (Grob et al., 2008). Zusätzlich sind vom System eventuelle Handlungsbedarfe anzuzeigen. Außerdem sollten konkrete Maßnahmen hinsichtlich ihrer möglichen Auswirkungen überprüft werden können (Hahn & Zwerger, 2002).

Tabelle 1: Anforderungen an ein PPMS

Dem methodischen Vorgehen des QFD folgend, sind in einem nächsten Schritt die Relevanz der erhobenen Anforderungen zu bestimmen (Akao, 2004).

4.2 Gewichtung der Anforderungen

Der AHP beginnt mit der Definition eines Entscheidungsproblems und dem Aufbau einer Zielhierarchie, in der die Anforderungen in verständliche und bewertbare Kriterien zerlegt werden. Das definierte Oberziel ist dabei auf der obersten Hierarchieebene anzuordnen, während die Alternativen auf der untersten Ebene der Hierarchie platziert werden. Hierbei ist zu berücksichtigen, dass jede Ebene der AHP-Hierarchie von der jeweils darüber liegenden Ebene beeinflusst wird. Im vorliegenden Anwendungsfall besteht das zu definierende Oberziel in der Gewichtung der Anforderungen an ein PPMS zur Unterstützung einer ganzheitlichen WoC-Steuerung. Die dreistufige Zielhierarchie besitzt neben dem genannten Oberziel, drei Kriterien und insgesamt 17 Attribute (s. Kap. 4.1) auf der untersten Hierarchieebene. Die Abb. 1 zeigt die aufgestellte Zielhierarchie.

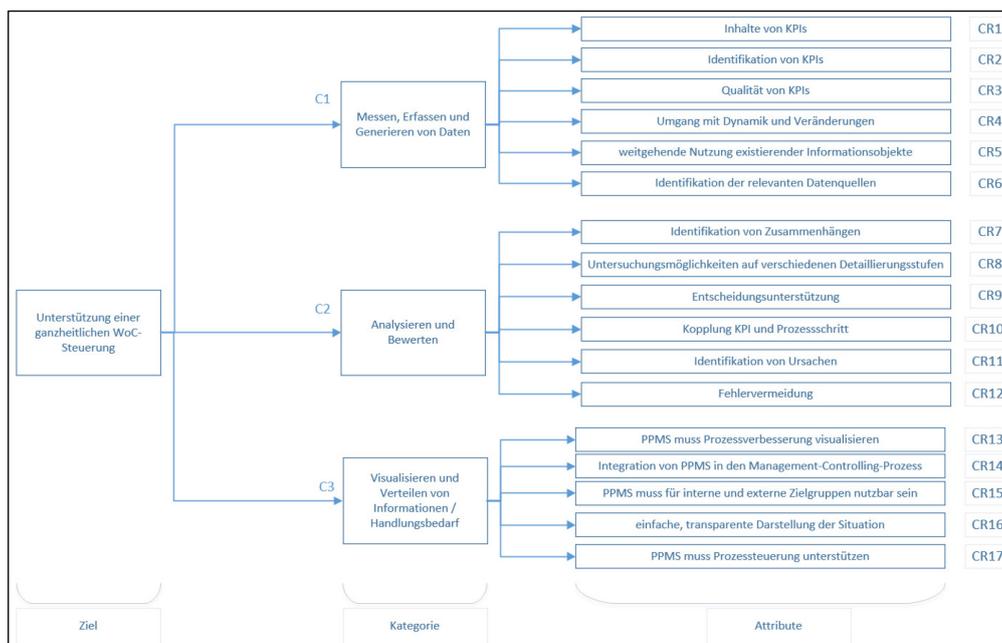


Abbildung 1: Zielhierarchie (eigene Darstellung)

Entsprechend der aufgestellten Hierarchie schließt sich die Gewichtung der einzelnen Anforderungen über einen paarweisen Vergleich der einzelnen Kriterien anhand einer metrischen Bewertungsskala an (Saaty, 2004). Auf Basis der erhobenen Präferenzurteile sind Vergleichsmatrizen zu bilden und die Gewichtungsvektoren auf Basis der Eigenvektor-

Methode zu berechnen (Saaty, 2003). Zur Gewichtung der Anforderungen in diesem Beitrag wurden die Prioritäten der Attribute, d.h. die Präferenzurteile auf Grundlage eines Expertenworkshops gebildet. Die befragte Zielgruppe wurde in diesem Zusammenhang sowohl mit den zu bewertenden Anforderungen als auch mit den methodischen Grundlagen des AHP vertraut gemacht, um die Qualität der Beurteilung zu stärken. Auf Grundlage der durch die Präferenzurteile ermittelten metrischen Werte werden in Abhängigkeit von der Struktur der Zielhierarchie vier Vergleichsmatrizen (C1, C2, C3, G) aufgebaut. Durch die anschließenden Matrizenberechnungen erfolgt die Ermittlung der Prioritäten ermittelt. Abbildung 2 zeigt beispielhaft die Vergleichsmatrix C1.

	CR1	CR2	CR3	CR4	CR5	CR6
CR1		3	1	3	6	3
CR2	0,333		0,333	3	3	3
CR3	1	3		3	3	6
CR4	0,333	0,333	0,333		3	3
CR5	0,167	0,333	0,333	0,333		0,333
CR6	0,333	0,333	0,167	0,333	3	

Abbildung 2: Vergleichsmatrix C1 (eigene Darstellung)

Die anschließende Eigenvektorberechnung ergibt folgende Gewichtungsvektoren:

$$\omega_{c1} = \begin{bmatrix} CR1 \\ CR2 \\ CR3 \\ CR4 \\ CR5 \\ CR6 \end{bmatrix} = \begin{bmatrix} 0,298 \\ 0,163 \\ 0,308 \\ 0,112 \\ 0,049 \\ 0,071 \end{bmatrix} \quad \omega_{c2} = \begin{bmatrix} CR7 \\ CR8 \\ CR9 \\ CR10 \\ CR11 \\ CR12 \end{bmatrix} = \begin{bmatrix} 0,054 \\ 0,146 \\ 0,427 \\ 0,077 \\ 0,053 \\ 0,243 \end{bmatrix} \quad \omega_{c3} = \begin{bmatrix} CR13 \\ CR14 \\ CR15 \\ CR16 \\ CR17 \end{bmatrix} = \begin{bmatrix} 0,104 \\ 0,145 \\ 0,049 \\ 0,609 \\ 0,092 \end{bmatrix}$$

$$\omega_G = \begin{bmatrix} C1 \\ C2 \\ C3 \end{bmatrix} = \begin{bmatrix} 0,655 \\ 0,250 \\ 0,095 \end{bmatrix}$$

Im nächsten Schritt muss eine Konsistenzprüfung durchgeführt werden, um die Validität der Ergebnisse zu bewerten. Die Konsistenz wird üblicherweise anhand der Faktoren „consistency ratio“ und „consistency measure“ bewertet, wobei ein consistency ratio von 0,2 als valide akzeptiert werden kann. Saraj 2011 führt weiterhin die Kongruenz und die Dissonanz als zusätzliche Größen an, um Ausreißer sowie einen „consistency deadlock“, d.h. eine völlige Inkonsistenz der Bewertungen, zu detektieren (Siraj, Mikhailov, & Keane, 2015).

matrix	consistency ratio	consistency measure	congruence	dissonance
C1	0,08	0,833	0,810	0,017
C2	0,054	0,667	0,630	0,0
C3	0,08	0,778	0,781	0,033
G	0,016	0,333	0,405	0

Tabelle 2: Konsistenzbetrachtung

Tabelle 2 zeigt, dass die Ergebnisse des paarweisen Vergleichs weitgehend konsistent sind. Die Ergebnisse liegen zum Teil sogar deutlich unterhalb des akzeptablen Werts von 0,2 (Mühlbacher & Kaczynski, 2013), was für eine gute Qualität der Präferenzurteile spricht.

Damit im Ergebnis eine kundenorientierte Priorisierung der Anforderungen an ein PPMS zur Unterstützung einer ganzheitlichen WoC-Steuerung erzielt wird, schließt sich die Berechnung der Gesamthierarchie aus den jeweiligen Teilgewichten als letzter Schritt im AHP an. Tabelle 3 zeigt die berechnete Gesamthierarchie und damit die relativen Bedeutungen der Anforderungen.

Ziel	Kategorie	Attribut
Kundenanforderungsgewichtung (1,00)	Messen, Erfassen und Generieren von Daten (0,655)	Inhalte von KPIs (0,195)
		Identifikation von KPIs (0,107)
		Qualität von KPIs (0,202)
		Umgang mit Dynamik und Veränderungen (0,073)
		weitgehende Nutzung existierender Informationsobjekte (0,032)
		Identifikation der relevanten Datenquellen (0,047)
	Analysieren und Bewerten (0,250)	Identifikation von Zusammenhängen (0,014)
		Untersuchungsmöglichkeiten auf verschiedenen Detaillierungsstufen (0,37)
		Entscheidungsunterstützung (0,107)
		Kopplung KPI und Prozessschritt (0,019)
		Identifikation von Ursachen (0,013)
	Visualisieren und Verteilen von Informationen / Handlungsbedarf (0,095)	Fehlervermeidung (0,061)
		PPMS muss Prozessverbesserung visualisieren (0,010)
		Integration von PPMS in den Management-Controlling-Prozess (0,014)
		PPMS muss für interne und externe Zielgruppen nutzbar sein (0,005)
	einfache, transparente Darstellung der Situation (0,058)	
	PPMS muss Prozesssteuerung unterstützen (0,009)	

Tabelle 3: Gesamthierarchie

Aus den Ergebnissen ist ersichtlich, dass der Kategorie „Messen, Erfassen und Generieren von Daten“ die größte Bedeutung zukommt. Bezogen auf einzelne Attribute ist bei der

Konstruktion einer kundenorientierten Lösung v.a. darauf zu achten, Untersuchungsmöglichkeiten auf verschiedenen Detaillierungsstufen zu implementieren sowie die Qualität von KPIs sicherzustellen. Die Ergebnisse unterstreichen auch die einleitend beschriebene Feststellung, dass es aktuell kein System für eine ganzheitliche WoC-Steuerung gibt, da die Priorisierung der Anforderungen vor allem auf der Schaffung relevanter Grundlagen, wie Qualität und Inhalten von KPIs liegt. Dagegen spielt eine geeignete Visualisierung dieser Kennzahlen eine untergeordnete Rolle.

5 Zusammenfassung und Ausblick

Im vorliegenden Beitrag wurden Anforderungen an ein PPMS zur Unterstützung einer ganzheitlichen WoC-Steuerung im Sinne einer kundenorientierten Priorisierung gewichtet. In diesem Zusammenhang wurden Grundlagen des WCM betrachtet. Die Komplexität einer ganzheitlichen Steuerung der zugrundeliegenden Prozesse zieht die Notwendigkeit eines geeigneten IT-Instrumentariums nach sich. Als einen möglichen Lösungsansatz schlugen die Autoren ein PPMS vor, dessen zuvor erhobenen Anforderungen näher beschrieben wurden. Die Gewichtung der Anforderungen im Kontext der ganzheitlichen WoC-Steuerung mittels der Methode des AHP sowie eines Expertenworkshops zeigte eine Priorisierung von grundlegenden Anforderungen, wie der Qualität und Inhalten von KPIs sowie detaillierten Untersuchungsmöglichkeiten dieser KPIs. Im Fokus steht damit vor allem das Messen, Erfassen und Generieren relevanter Daten. Diese Ergebnisse fließen in die weitere theoretische Konzeption und prototypische Implementierung des PPMS zur Unterstützung einer ganzheitlichen WoC-Steuerung im Rahmen der weiteren Forschungstätigkeiten der Autoren ein.

Literaturverzeichnis

- Akao, Y. (2004). *Quality Function Deployment*: Productivity Press.
- Aktas, N., Croci, E., & Petmezas, D. (2015). Is working capital management value-enhancing? Evidence from firm performance and investments. *Journal of Corporate Finance*, 30, 98–113.
- Buchheim, R. K. (2000). Developing performance metrics for a design engineering department. *IEEE Transactions on engineering management*, 47(3), 309–320.
- Cleven, A. (2012). Process Performance Measurement - a systematic problem analysis and identification of design principles. *Information Systems Journal of Organisational Design and Engineering*. (3), 227–249.
- del-Río-Ortega, A., Resinas, M., Cabanillas, C., & Ruiz-Cortés, A. (2013). On the definition and design-time analysis of process performance indicators. *Special section on BPM 2011 conference*, 38(4), 470–490.
- Döring, O., Schönherr, M., & Steinhäuser, P. (2012). Working Capital Controlling: Entwicklung eines Steuerungskonzepts zur nachhaltigen Performancesteigerung eines mittelständischen Industrieunternehmens. *Controlling*. (8/9), 409–415.

- Enqvist, J., Graham, M., & Nikkinen, J. (2014). The impact of working capital management on firm profitability in different business cycles: Evidence from Finland. *Research in International Business and Finance*, 32, 36–49.
- Gladen, W. (2002). Performance Measurement als Methode der Unternehmenssteuerung. *HMD-Praxis der Wirtschaftsinformatik*, (227), 5–16.
- Glavan, L. (2011). Understanding process performance measurement systems. *Business Systems Research*, 2(2), 25–38.
- Grob, L. H., Bensberg, F., & Coners, A. (2008). Rule-based Control of Business Processes – A Process Mining Approach. *Wirtschaftsinformatik*, 50(4), 268–281.
- Grosswiele, L., Röglinger, M., & Friedl, B. (2013). A decision framework for the consolidation of performance measurement systems. *Decision Support Systems*, 54(2), 1016–1029.
- Hahn, A., & Zwerger, G. (2002). Performance Measurement Software Tools. *IEEE Transactions on engineering management*, 227, 97–102.
- Hofmann, E., & Martin, J. (2014). *Working Capital Management-Performance Excellence-Studie Band 2014: St.Galler WCM-Studie* (1st ed.). Göttingen: Cuvillier. Retrieved from <https://www.alexandria.unisg.ch/publications/237393>
- Klepzig, H.-J., & Vater, H. (2015). Fat Cash durch Lean Management - der Treiber: Working Capital. *Controller Magazin*, (4), 62–67. Retrieved from <http://zeitschriften.haufe.de/ePaper/controller-magazin/2015/C1A5AE95/files/assets/basic-html/page64.html>
- Knauer, T., & Wöhrmann, A. (2013). Working capital management and firm profitability. *Journal of Management Control*, 24(1), 77–87.
- Kudernatsch, D. (2002). Performance Measurement im IT-Management. *Praxis der Wirtschaftsinformatik: HMD*, 227, 56–66.
- Kueng, P. (2000). Process performance measurement system: a tool to support process-based organizations. *Total Quality Management*, 11(1), 67–85.
- Kueng, P., & Krahn, A. J. W. (1999). Building a process performance measurement system: some early experiences. *Journal of scientific and industrial research*, 58, 149–159.
- Lind, L., Pirttilä, M., Viskari, S., Schupp, F., & Kärrä, T. (2012). Working capital management in the automotive industry: Financial value chain analysis. *Journal of Purchasing and Supply Management*, 18(2), 92–100.
- Meixner, O., & Haas, R. (2002). *Computergestützte Entscheidungsfindung: expert choice und AHP-innovative Werkzeuge zur Lösung komplexer Probleme*: Redline Wirtschaft bei Ueberreuter Frankfurt.
- Mühlbacher, A. C., & Kaczynski, A. (2013). Der Analytic Hierarchy Process (AHP): Eine Methode zur Entscheidungsunterstützung im Gesundheitswesen. *PharmacoEconomics German Research Articles*, 11(2), 119–132.
- Pfützner, M., & Hilbert, A. (2014). Anforderungen an ein Analytisches Working Capital Management. *Corporate Finance: Finanzierung, Kapitalmarkt, Bewertung, Mergers & Acquisitions*, 1(5), 33–39.
- Protopappa-Sieke, M., & Schimke, M. (2010). Working Capital Management in der Supply Chain: Es geht ums Geld. *BA Beschaffung aktuell*, (3), 22–24.
- Robson, I. (2004). From process measurement to performance improvement. *Business Process Management Journal*, 10(5), 510–521.

- Rupp, R. (2011). Working Capital Management-Controlling mit eindrucksvollen Bildern oder mit belastbaren Zahlen. *Controlling-Zeitschrift für erfolgsorientierte Unternehmenssteuerung*, 23, 379–386.
- Saaty, T. L. (2003). Decision-making with the AHP: Why is the principal eigenvector necessary. *European Journal of Operational Research*, 145(1), 85–91.
- Saaty, T. L. (2004). Decision making — the Analytic Hierarchy and Network Processes (AHP/ANP). *Journal of Systems Science and Systems Engineering*, 13(1), 1–35.
- Siraj, S., Mikhailov, L., & Keane, J. A. (2015). PriEsT: an interactive decision support tool to estimate priorities from pairwise comparison judgments. *International Transactions in Operational Research*, 22(2), 217–235.
- Sure, M. (2015). Working Capital-Steuerung mit integrierten Kennzahlen. *Controlling*, 27(1), 7–12.
- Wettstein, T., Kueng, P., & Meier, A. (2001). Performance Measurement als Ausweg aus dem Information Overload: Ein zielorientierter Ansatz. *HMD-Praxis der Wirtschaftsinformatik*, 222, 49–58.
- Wieland, U., Pfitzner, M., Fischer, M., & Hilbert, A. (2015). Process performance measurement system – towards a customer-oriented solution. *Business Process Management Journal*, 21(2), 312–331.
- Xirogiannis, G., & Glykas, M. (2004). Fuzzy cognitive maps in business analysis and performance-driven change. *IEEE Transactions on engineering management*, 51(3), 334–351.

Kontaktinformationen

Dipl.-Kfm. Marcus Pfitzner
Technische Universität Dresden
Lehrstuhl für Wirtschaftsinformatik | Business Intelligence Research
Münchner Platz 3
D-01187 Dresden

Tel.: +49 351 463 33629

E-Mail: <mailto:marcus.pfitzner@tu-dresden.de>

WWW <https://tu->

[dresden.de/die_tu_dresden/fakultaeten/fakultaet_wirtschaftswissenschaften/wi/wiid](https://tu-dresden.de/die_tu_dresden/fakultaeten/fakultaet_wirtschaftswissenschaften/wi/wiid)

Promotionsvorhaben zur Erweiterung des Algorithmus CLTree

Martin Thrum¹

Lehrstuhl Wirtschaftsinformatik, Friedrich-Schiller-Universität Jena¹

Zusammenfassung

Eine Beobachtung im Bereich der Data Mining Algorithmen ist, dass Verfahren in den primären Veröffentlichungen als erfolgreich dargestellt werden, anschließend aber keine Verbreitung erlangen. Der von (Liu et al., 2000) vorgestellte Algorithmus CLTree wurde als breit anwendbarer Algorithmus für verschiedene Clusteranwendungen publiziert. Neben den klassischen Clusteraufgaben ist er auch für Subspace-Clustering und das Auffinden von nicht besiedelten Bereichen in den Daten geeignet. Trotz der vielfältigen Anwendbarkeit, ist CLTree kaum bekannt und ebenso wenig in Software umgesetzt. In diesem Aufsatz wird ein Promotionsvorhaben vorgestellt, das Gründe für die geringe Verbreitung des Algorithmus durch ein theoretisches Review in Kombination mit ausführlichen Testfällen untersucht. Im Ergebnis soll neben der ausführlichen Bewertung der Leistungsfähigkeit eine Reihe an Erweiterungen entstehen, die eventuell gefundene Schwachstellen adressieren und beheben.

1 Einleitung

Aktuelle Trends der datengetriebenen Industrien und den zugrunde liegenden Forschungsrichtungen beschäftigen sich mit der Analyse großer Datenbestände.¹ Neben den Techniken zur Speicherung und dem Zugriff auf diese Daten spielen zur Generierung von Wissen auch Data Mining Algorithmen eine Rolle. Regelmäßig werden von Wissenschaftlern neue Algorithmen vorgeschlagen (Kumar & Bhardwaj, 2011). Die Vielzahl verschiedener Verfahren steht jedoch im Gegensatz zu den tatsächlich im Data Mining verwendeten Algorithmen. Die Liste der häufig verwendeten Algorithmen wird nach wie vor von etablierten Verfahren wie z.B. k-means, C4.5 oder k-NN angeführt (Wu & Kumar, 2009). Eine mögliche Ursache für diese Beobachtung ist neben den häufig fehlenden Implementierungen der neu

¹ z.B. <http://www.journals.elsevier.com/big-data-research> (Stand 28.09.2016)

vorgestellten Verfahren auch eine fehlende ausführliche, über den veröffentlichten Pirmäraufsatz hinausgehende Evaluation der Algorithmen. Der hier betrachtete Algorithmus CLTree wurde im Jahr 2000 veröffentlicht und seitdem in rund 200 wissenschaftlichen Aufsätzen referenziert.² Eine genaue Betrachtung ergibt allerdings, dass es sich hier häufig nur um Nennungen im „related work“ Bereich handelt und nur wenige Arbeiten, wie z.B. (Parsons et al., 2004) oder (Juozapavičius & Rapševičius, 2001), sich tatsächlich mit dem Verfahren auseinandersetzen. Zudem hat die eigene Recherche ergeben, dass bisher keine Implementierung in öffentlich verfügbaren Data Mining Programmen existiert.³

Die in der Erstveröffentlichung angegebenen Stärken und Einsatzmöglichkeiten des Verfahrens stehen den oben beschriebenen Beobachtungen gegenüber im Kontrast. Vor allem die breite Anwendbarkeit von CLTree auf verschiedene Clusteraufgaben wie z.B. das Aufdecken von Löchern in Daten oder die Eignung für Subspace-Clustering Analysen bieten gute Ansatzpunkte für weitere Untersuchungen. Als Ziele der Dissertation werden die ausführliche Evaluation des Algorithmus aus theoretischer als auch empirischer Sicht abgeleitet sowie die Erweiterung zur Lösung aufgedeckter Schwächen angestrebt. Der Aufbau gliedert sich dabei nach der Analyse der Ausführungen im Primärpaper in zwei Testblöcke. Die Schwerpunkte der theoretischen Untersuchung liegen auf der Überprüfung der vorgeschlagenen Verfahrensschritte sowie des Pseudo-Codes auf Konsistenz und Eindeutigkeit. Während der ersten Testreihe werden anschließend die zuvor gewonnenen Erkenntnisse in eine Referenz-Implementierung übernommen. Nicht eindeutig beschriebene Arbeitsschritte sind als Parameter umzusetzen und so das Verhalten des Algorithmus zu katalogisieren. Für die Tests werden neben verschiedenen synthetischen Datensätzen, die jeweils für spezielle Szenarien konstruiert sind, auch in der Forschung etablierte Datensätze wie z.B. der iris-Datensatz (Lichmann, 2013) zur Evaluation verwendet. Auf Grundlage der hier gewonnenen Erkenntnisse können Erweiterungen für den CLTree Algorithmus vorgeschlagen werden, die die entdeckten Schwächen adressieren und lösen sollen. Die Evaluation der genannten Erweiterungen erfolgt anhand der gleichen Datensätze, welche auch im ersten Testblock zur Anwendung kommen. Auf diese Weise wird die Vergleichbarkeit der Tests gewährleistet. In der abschließenden Auswertung der Beobachtungen erfolgt die Beschreibung geeigneter Anwendungsszenarien für den erweiterten CLTree Algorithmus.

2 Auswahl der Quellen zur Literaturrecherche

Die Auswahl der Quellen zur fachlichen Absicherung der Arbeit wird durch ein breites Spektrum an verwandten Themengebieten geprägt. Im Gegensatz reinen Aufgabenstellungen der Klassifikation oder Clustering von Datensätzen ist CLTree mit beiden

² Online: 33 Referenzen auf ACM digital library [<http://dl.acm.org/>] (Stand 29.09.2016) und 185 auf Google scholar [<https://scholar.google.de/>] (Stand 29.09.2016)

³ Überprüft in Weka 3.7, KNIME 3.2.1, Orange 3.3.7, RapidMiner 5.3, R 3.1.3

Forschungsbereichen verbunden. Die Schwerpunkte zur Recherche liegen dabei auf Literatur zu:

1. Entscheidungsbaumalgorithmen allgemein,
2. Splitkriterien v.a. in Hinblick auf vorausschauende Splitverfahren,
3. Abbruchkriterien beim Baufbau und Pruningstrategien,
4. Maßen zur Clustervalidierung und
5. dem Auffinden von Löchern in Datenbeständen sowie
6. Subspace-Clustering.

Während die ersten drei genannten Punkte häufig in gemeinsamen Quellen wie z.B. den Data Mining Sammelwerken (Han et al., 2011) oder (Tan et al., 2005) in einem Kontext behandelt werden, entstammt die Validierung von Clusterlösungen dem Forschungsbereich des unüberwachten Lernens. Die Recherche in diesem Fachgebiet ist vor allem interessant, um eventuell geeignete Maßzahlen zur aktiven Qualitätskontrolle in den Prozess des Baufbaus zu integrieren. Aus dem Fachgebiet der Suche nach leeren bzw. gering besiedelten hyperdimensionalen Bereichen in Datensätzen sollen Techniken entlehnt werden, die der Bestimmung eben dieser maximalen leeren Räume dienen. Im Umkehrschluss können mit diesen Verfahren ebenfalls Clustergrenzen ermittelt werden. Eine ähnliche Rolle wie die Suche nach Löchern spielt die Disziplin des Subspace-Clustering. Der Fachbereich wird als mögliches Anwendungsgebiet des Algorithmus genannt (Liu et al., 2000) und sollte dementsprechend mindestens zur Ermittlung geeigneter Anwendungsszenarien herangezogen werden.

Ein weiteres Themengebiet, das zu nennen und abzugrenzen ist, umfasst die Algorithmen des hierarchischen Clusters. Vor allem während der Suche nach Implementierungen des CLTree Algorithmus tauchen die hierarchischen Clusterverfahren wie z.B. AGNES (Kaufman & Rousseeuw, 1990) immer wieder mit auf. Das gleichartige Vokabular zur Beschreibung der entstehenden Modellstruktur „Baum“ in Kombination mit der Anwendung „Clustern“ begünstigt aufgrund des höheren Bekanntheitsgrades den Vorschlag von Veröffentlichungen, die nicht mit CLTree in Zusammenhang stehen. Eine saubere Abgrenzung zwischen den Gebieten ist sehr wichtig, da die erstellten Modelle grundlegend verschieden sind. Eine Abhandlung zur Einteilung von Algorithmen nach ihren Prinzipien zur Modellbildung beschreibt die Notwendigkeit der Abgrenzung ausführlich (Estivill-Castro, 2002).

3 Der Algorithmus CLTree

Zum Verständnis der weiteren Ausführungen und Tests folgt eine kurz gefasste Beschreibung des Algorithmus CLTree nach (Liu et al., 2000). Die algorithmische Grundlage bildet das Verfahren C4.5 (Quinlan, 1993). Dieser Baumalgorithmus entscheidet anhand des Informationsgewinn-Splitkriteriums, an welcher Stelle jeder Dimension d_i über alle verfügbaren Dimensionen D eine optimale Aufteilung des betrachteten Datenbestands erfolgen kann. Aus allen ermittelten möglichen Ergebnissen werden die Daten dann entlang derjenigen Dimension unterteilt, bei der die mittlere Entropie-Differenz maximal wird. Da zur

Berechnung dieser Reinheit Informationen zu den Klassen notwendig sind, ist der Algorithmus per se nicht zur Clusterung von Daten verwendbar.

Um diese Herausforderung zu bewältigen schlagen (Liu et al., 2000) vor, die tatsächlich existierenden Datenpunkte der Klasse der Y-Punkte zuzuordnen und eine zweite Klasse an imaginären N-Punkten einzuführen. Diese N-Punkte liegen gleichverteilt im gesamten Datenraum vor. Durch diese Änderung kann das Clusterproblem in ein Klassifikationsproblem umgewandelt werden. Der Entscheidungsbaum besitzt am Ende der Induktionsphase Blattknoten, die vorwiegend Y-Punkte enthalten, also die Cluster darstellen und Blattknoten, die vorwiegend N-Punkte enthalten, also die leeren Bereiche im Datensatz beschreiben. Der Baumaufbau endet, sobald keine Unterteilung der Teilmengen mehr möglich ist.

Zur Wahrung der Performance von Entscheidungsbaumalgorithmen werden die N-Punkte nicht tatsächlich im Datensatz eingefügt. Stattdessen erfolgt die Berechnung der pro Teilmenge vorhandenen N-Punkte online während der Splitberechnung relativ zum aktuell betrachteten Splitpunkt anhand des verfügbaren Wertebereichs.

Eine zweite wesentliche Änderung am ursprünglichen C4.5 Algorithmus schlagen (Liu et al., 2000) am Splitkriterium an sich vor. Das standardmäßig verwendete Informationsgewinnkriterium, welches den Entropieverlust der Teilmengen durch einen Split lokal maximiert, tendiert nach ihrer Aussage dazu, in Clusterstrukturen hineinzuschneiden. Als mögliche Lösung wird ein vorausschauendes Splitkriterium empfohlen, welches entlang jeder Dimension weitere Splits auf Basis des zuerst gefundenen sucht. Ergeben sich durch das vorausschauende Unterteilen Teilmengen, die ein besseres Y-N-Punkte-Verhältnis aufweisen als die vorhergehenden, so wird der dazugehörige Split verwendet.

4 Review-Ergebnisse der Algorithmenbeschreibung

Da eine vollständige Implementierung für Algorithmen nur selten zur Darstellung im Fließtext geeignet ist, werden in Veröffentlichungen abgekürzte Beschreibungen in Pseudo-Code aufgeführt. Diese notwendige Vereinfachung birgt jedoch die Gefahr, dass nicht alle Entscheidungen zum Ablauf des Verfahrens mit eingeschlossen sind. Wenn derartige Informationslücken nicht durch eine Veröffentlichung des Quellcodes oder durch separate verbale Ausführungen gefüllt werden, ist die Umsetzung des Algorithmus durch Dritte nicht mehr eindeutig möglich.

Die Darstellung von CLTree nach (Liu et al., 2000) besteht aus drei separat vorgestellten Erweiterungen. Diese sind die Aktualisierung der N-Punkte, die Auswahl des besten Splits entlang eines Attributs unter Verwendung des vorausschauenden Splitkriteriums und das Verfahren zum nutzerorientierten Pruning des Baums. Eine vollständige Formulierung des Algorithmus ist nicht vorhanden. Da der grundlegende Ablauf auf dem bekannten C4.5 Algorithmus beruht und nur die Veränderungen dargestellt werden müssen, ist diese Entscheidung nachvollziehbar.

Bei der eigenen Implementierung und dem damit verbundenen Review des Algorithmenablaufs an sich sind bisher zwei Informationslücken aufgefallen, die es in der ersten Testreihe zu analysieren gilt.

1. Bedingung der Ganzzahligkeit der N-Punkte
2. Aktualisierung der N-Punkte während eines vorausschauenden Splits

Auf die Notwendigkeit der Ganzzahligkeit der N-Punkte gehen (Liu et al., 2000) nicht ein. In den gewählten Beispielen ergeben sich Splits immer an Stellen, bei denen die N-Punkte ganzzahlig aufgeteilt werden können. Die Forderung nach ganzzahligen N-Punkten ist einerseits plausibel, da die Y-Punkte auch atomar sind, andererseits handelt es sich bei N-Punkten um imaginäre Konstrukte, die dieser Restriktion nicht unterliegen. Ein Parameter für das Hinzufügen oder Abwählen dieser Bedingung wird in die erste Testreihe mit aufgenommen. Es ist zu erwarten, dass, wenn sich überhaupt ein messbarer Unterschied ergibt, dieser erst bei sehr großen Datensätzen mit komplex geformten Clustern auftritt.

Zur Aktualisierung der N-Punkte während eines vorausschauenden Splitvorgangs hingegen kann noch keine direkte Abschätzung über ein Testergebnis abgegeben werden. Die hier bezeichnete Problematik adressiert die Formulierung, dass bei jedem neuen Split die N-Punkte der Kindknoten an die der Elternknoten anzupassen sind. Allerdings erklären (Liu et al., 2000) nicht, ob diese Überprüfung auch während eines vorausschauenden Splits entlang einer Dimension durchzuführen ist.

5 Aufbau des Testframeworks – Daten und Parameter

Da die in der Veröffentlichung von (Liu et al., 2000) dargestellten Testergebnisse mangels öffentlich zugänglicher Implementierung und Datensätze nicht nachvollziehbar sind, kann nur in Annäherung an die Beschreibungen im Paper ein Vergleich der selbst erstellten Implementierung mit dem Original erfolgen.

Der Testaufbau an sich wird neben den zu bestimmenden Parametern⁴ vor allem durch die Charakteristik der verwendeten Datensätze bestimmt. Die Orientierung zur Auswahl der Daten geben v.a. Experimente aus dem Gebiet der Entscheidungsbäume und dem Auffinden von Löchern in Daten. Zudem werden Datensätze mit bestimmten Eigenschaften konstruiert. Diese werden durch verschiedene, den Clustern zugrunde liegende Verteilungen, damit verbunden den Clusterformen und der Aufteilung auf verschiedene Dimensionen gekennzeichnet. Die folgende Auflistung umfasst die wesentlichen Parameter, nach denen die Datensätze aufgebaut werden:

1. Clusterformen, –dichten und Anzahl
2. Verteilungen der Daten im Cluster
3. Anzahl der Dimensionen
4. Vorhandensein und Intensität von Rauschen in den Daten

⁴ z.B. maximale Baumtiefe, alternative Splitmaße wie Gini-Index, etc.

5. Verwendung unterschiedlicher Datentypen

Bei bestehenden Datensätzen wie z.B. dem häufig verwendeten Iris- oder Sojabohnen-Datensatz (Lichman, 2013) wird darauf geachtet, dass entweder vergleichbare Lösungen mit bekannter Güte oder Klassenlabel für die Daten an sich existieren. Diese Informationen sollen eine spätere Evaluation der gefundenen Lösungen ermöglichen. Die Auswahl einer Maßzahl zur Evaluation der Lösungen ist noch nicht abgeschlossen. Als bekannte Möglichkeiten stehen das F-Measure (Zhang & Zhang, 2009) bei Vorhandensein von Klassenlabeln oder der Silhouette-Koeffizient (Rousseeuw, 1987) für Daten auf denen eine Distanzberechnung möglich ist zur Verfügung. Beide Verfahren sind allerdings nicht universell für alle Datensätze einsetzbar, sodass aktuell noch nach einer geeigneten finalen Bewertungsmaßzahl recherchiert wird. Als Option ist hierbei auch die Konstruktion einer verfahrensspezifischen neuen Maßzahl denkbar, die jedoch ebenfalls selbst ausführlich zu bewerten ist.

Abschließend sind Tests mit Datensätzen aus einem oder mehreren, in der Anwendungsfallanalyse ermittelten, Szenarien geplant. Da diese Anwendungsfälle noch nicht ausgearbeitet wurden, konnten noch keine passenden Datensätze erhoben werden.

6 Ausblick auf mögliche Erweiterungen

Während der Erarbeitung des Themas erfolgte bereits ein prototypischer Durchlauf des bisher beschriebenen Testplans. Auf diese Weise wurde festgestellt, ob Verbesserungen für den Algorithmus gefunden werden können. Nach wenigen Experimenten haben sich bereits zwei Probleme sowie denkbare Lösungen dafür ergeben. Beide Probleme treten bei der Ermittlung des optimalen Splitpunkts auf.

Eine Auffälligkeit weist CLTree in der Wahl der Dimension auf, die den optimalen Split bietet. Es wurde beobachtet, dass Splits häufig entlang der ersten Dimension durchgeführt werden, obwohl in anderen Dimensionen größere Bereiche des Datensatzes abgetrennt werden können. Die Analyse hat ergeben, dass die Verwendung der relativen Dichte, berechnet als das Verhältnis der Y-Punkte zu den N-Punkten, als entscheidende Maßzahl häufig zu gleichwertigen optimalen Werten für verschiedene Dimensionen gelangt. Begründet liegt dies im vorausschauenden Splitverfahren, welches für nahezu jede Verteilung der Datenpunkte einen Bereich mit einer relativen Dichte von null ermittelt. Eine Handlungsanweisung für den Fall, dass mehrere Dimensionen diese optimale relative Dichte erreichen, wird von (Liu et al., 2000) nicht zur Verfügung gestellt. Damit ergeben sich Splitfolgen, wie sie in *Abbildung 5* links gekennzeichnet sind. Durch die Auswahl des ersten Attributs, dessen Split eine relative Dichte von null aufweist, werden die eventuell besser geeigneten Splits späterer Dimensionen nicht mit berücksichtigt. Neben der einfachsten, aber wenig robusten Lösung – der Zufallsauswahl – bietet sich die zusätzliche Betrachtung des Informationsgewinnkriteriums zur Entscheidungsfindung an. Mit dieser Erweiterung war es möglich die Tiefe der aufgebauten Bäume zu verringern und Splits zielgerichteter auf Dimensionen zu verteilen (siehe *Abbildung 5* rechts).

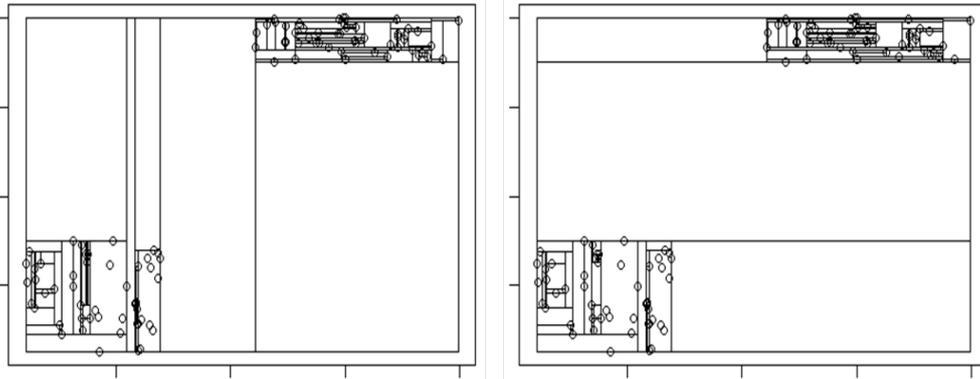


Abbildung 5: Splitverhalten unter Berücksichtigung des Informationsgewinns (links: ohne, rechts: mit)

Die zweite bisher entdeckte Problematik besteht in dem zwingenden Durchführen des vorausschauenden Splits. Für Dimensionen mit einer hohen gleichbleibenden Datendichte, allerdings mit einem oder mehreren eindeutig optimalen Punkten für Splits nach dem Informationsgewinnkriterium ergeben sich eigenwillige Splitpunkte wie der in *Abbildung 6* dargestellte. Der Grund hierfür ist die Suche nach einem weiteren Split entlang der Dimension, der die relative Dichte minimiert und den Informationsgewinn wieder außen vor lässt. Als mögliche Lösungen für dieses Problem ist ein Toleranzbereich für vorausschauende Splits in Hinblick auf den Informationsgewinn denkbar. Wird dieser Toleranzbereich bei der weiteren Suche nach Splitpunkten verlassen, so bricht das Verfahren ab und schlägt nur den zuerst gefundenen Punkt als mögliches Optimum vor. Bei diesem Vorschlag handelt es sich bisher nur um eine erste Idee. Eine eingehende Untersuchung ist noch durchzuführen.

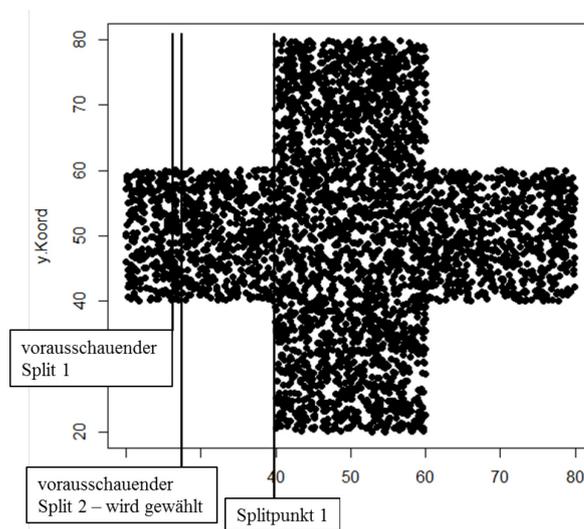


Abbildung 6: Problematik ungünstiger Splitauswahl bei dichten Daten

Literaturverzeichnis

- Estivill-Castro, V. (2002). *Why so many clustering algorithms: a position paper*. SIGKDD Explor. Newsl., 65-75.
- Han, J., Kamber, M., & Pei, J. (2011). *Data Mining: Concepts and Techniques*. Burlington: Elsevier Science.
- Juozapavičius, A., & Rapševičius, V. (2001). *Clustering through Decision Tree Construction in Geology*. In: Vilnius University (Hrsg.): *Nonlinear Analysis: Modelling and Control*. v.6 No. 2. Vilnius, Lithuania. 29-41.
- Lichman, M. (2013). UCI Machine Learning Repository [<http://archive.ics.uci.edu/ml>]. Irvine, CA: University of California, School of Information and Computer Science.
- Liu, B., Xia, Y., & Yu, P.S. (2000). *Clustering Through Decision Tree Construction*. McLean, Virginia. ACM CIKM-2000.
- Kaufman, L., & Rousseeuw, P.J. (1990). *Finding Groups in Data: An Introduction to Cluster Analysis*. Wiley: New York.
- Kumar, D., & Bhardwaj, D. (2011). *Rise of Data Mining: Current and Future Application Areas*. IJSCI Vol. 8, Issue 5, No 1, 256-260.
- Parsons, L., Haque, E., & Liu, H. (2004). *Evaluating Subspace Clustering Algorithms*. Arizona State University.
- Quinlan, J.R. (1993). *C4.5 programs for machine learning*. San Mateo, California. Kaufmann.
- Rousseeuw, P. (1987). *Silhouettes: a graphical aid to the interpretation and validation of cluster analysis*. Journal of Computational and Applied Mathematics. No. 20. North-Holland. 53-65.
- Tan, P., Steinbach, M., & Kumar, V. (2005). *Introduction to Data Mining*. Boston: Pearson Education, Inc.
- Wu, X., & Kumar, V. (2009). *The Top Ten Algorithms in Data Mining*. CRC Press.
- Zhang, E., & Zhang, Y. (2009). *Encyclopedia of Database Systems*. Springer.

Kontaktinformationen

M.Sc. Martin Thrum
Friedrich-Schiller-Universität Jena
Lehrstuhl für Wirtschaftsinformatik
Carl-Zeiss-Str. 3
D-07743 Jena

Tel.: +49 (0)3641 9-43319

E-Mail: martin.thrum@uni-jena.de

WWW <http://www.wiinf.uni-jena.de/>

Unternehmensübergreifende Stammdatenqualität – Entwicklung eines Hilfsmittels zur Vereinbarung der Qualität für Stammdaten zwischen Unternehmen

Thomas Schäffer¹

Fakultät für Wirtschaftswissenschaften und Medien im Fachgebiet Informations- und
 Wissensmanagement, Technische Universität Ilmenau¹

Studiengang Wirtschaftsinformatik, Hochschule Heilbronn¹

Zusammenfassung

Stammdaten bilden die Grundlage der digitalen Wirtschaft. Die Bereitstellung eines angemessenen Maßes an Stammdatenqualität ist eine entscheidende Voraussetzung für effiziente Geschäftsprozesse in und zwischen Unternehmen. Der Beitrag stellt das Promotionsvorhaben vor, welches sich mit der Entwicklung eines Hilfsmittels beschäftigt, mit dem zwei Unternehmen eine Vereinbarung und Prüfung des Qualitätsniveaus für Stammdaten, die zwischen den beiden Unternehmen übertragen werden, weitgehend automatisiert durchführen können. Im Rahmen des Dissertationsvorhabens durchgeführte Analysen zeigen, dass in vielen Unternehmen der Produktstammdatenaustausch aufgrund mangelhafter Stammdatenqualität problembehaftet, fehleranfällig, arbeits- und kostenintensiv ist. Der Einsatz von Basistechnologien und Standards für den Stammdatenaustausch reicht meist nicht aus, um die spezifischen Anforderungen an die Stammdatenqualität zu spezifizieren und zudem sind die heutigen Werkzeuge zu kostenintensiv und nicht flexibel genug in den Augen vieler Experten in Unternehmen.

1 Einleitung

Erhöhte Digitalisierung in Wirtschaft und Gesellschaft erfordert drastische Veränderungen in den Unternehmen. Laut BDI-Studie (Bloching et al., 2015) wird die Unternehmensvernetzung stetig wachsen und sich eine Verschiebung von starren Wertschöpfungsketten hin zu dynamischen Wertschöpfungsnetzwerken einstellen. Stammdaten bilden die Grundlage der digitalen Wirtschaft und sind für den Informationsaustausch innerhalb von Unternehmensnetzwerken von essentieller Bedeutung (GS1 Germany, 2016; Otto & Österle, 2016). Das Erreichen und die Sicherstellung eines angemessenen Maßes an Stammdatenqualität ist eine kritische Voraussetzung für eine effiziente und effektive unternehmensübergreifende Zusammenarbeit (Dalmolen et al., 2015; de Corbière, 2007; Hüner, et al., 2011; Legner & Wende, 2007; Loshin, 2008; Otto et al., 2014).

Stammdaten beschreiben kritische Geschäftsobjekte eines Unternehmens (Loshin, 2008) und bezeichnen Produkte, Lieferanten, Kunden, Mitarbeiter bzw. ähnliche Gegenstände, die nur selten Änderungen erfahren (Haug & Stentoft Arlbjörn, 2011; Ofner et al., 2013). Die Sicherstellung und eines angemessenen Qualitätsniveau der Stammdatenpflege ist besonders wichtig (Wang, 1998).

Datenqualität ist "ein Maß für die Eignung der Daten für spezifische Anforderungen in Geschäftsprozessen, in denen sie verwendet werden. Die Datenqualität ist ein mehrdimensionales, kontextuelle Konzept, da es nicht mit einer einzigen Funktion beschrieben werden kann, sondern auf der Basis verschiedener Datenqualitätsdimensionen und Metriken" (Otto & Österle, 2016). Typische, häufig verwendete Qualitätsdimensionen sind Fehlerfreiheit, Korrektheit, Vollständigkeit, Relevanz, Konsistenz und Aktualität. Die Datenqualität wird deshalb oft mit dem Begriff "fitness for use" assoziiert (de Corbière, 2007; Wang & Strong, 1996).

Unternehmensübergreifende Geschäftsprozesse bestehen aus logisch zusammenhängende Aktivitäten, die sich über mehr als einzelnes Unternehmen erstrecken (Bakos, 1991) und zum Erreichen eines definierten Geschäftsergebnis durchgeführt werden (Davenport & Short, 1990). Eine typische Aufgabe in einem unternehmensübergreifenden Geschäftsprozess ist der Produktstammdatenaustausch zwischen Unternehmen.

Produktstammdaten enthalten eine Anzahl von Merkmalen, beispielsweise Identifikationsnummer, Gewicht, Größe usw., die das Produkt beschreiben (Dalmolen et al., 2015; de Corbière, 2007; Madlberger, 2011). Produktstammdatenaustausch bezeichnet die unternehmensübergreifende Übertragung von Produktstammdaten und in diesem Zusammenhang verwenden Legner und Schemm (2008) die Bezeichnung "Produkt Information Supply Chain".

Die wissenschaftliche Literatur beschäftigte sich in der Vergangenheit mit zahlreichen Beiträgen zur Sicherstellung und Pflege der unternehmensinternen Stammdatenqualität (Otto & Österle, 2016; Wand & Wang, 1996; Wang & Strong, 1996). Im Gegensatz dazu wurde die Stammdatenqualität in unternehmensübergreifenden Geschäftsprozessen wenig bis gar nicht

beleuchtet (Dalmolen et al., 2015; de Corbière, 2007; Legner & Schemm, 2008; Madlberger, 2011).

Sowohl aus der wissenschaftlichen als auch aus der praktischen Perspektive ist zu erkennen, dass die Stammdatenübermittlung, das heißt, unternehmensübergreifende Übertragung von Stammdaten, ein problematischer, fehleranfälliger, arbeitsintensiver und kostspieliger Prozess in vielen Unternehmen ist (Madlberger, 2011; Ofner et al., 2013).

Zahlreiche wissenschaftliche Publikationen identifizierten den Forschungsbedarf zur Vereinbarung der Qualität für Stammdaten zwischen Unternehmen (Haug & Stentoft Arlbjörn, 2011; Legner & Wende, 2007; Loshin, 2008; Ofner, Otto, & Österle, 2012). Bei der Auseinandersetzung mit dieser Forschungslücke, stellten mehrere Autoren fest, dass beispielsweise Normen zur Produktidentifikation und Klassifizierung und Datenpools keine breite Akzeptanz haben, weil sie wesentliche spezifische Anforderungen von Unternehmen nicht erfüllen (Bakos, 1991; Dalmolen et al., 2015; Madlberger, 2011; Schemm, 2009). Auch für den Datenaustausch per EDI (z. B. PRICAT) wird darauf hingewiesen, dass zwischen dem Sender und dem Empfänger der EDI-Nachrichten ein höheres Maß an Kontext erforderlich ist und allgemein Standards für die Erstellung von System-zu-System-Integrationen alleine nicht ausreichend sind (Helmis & Hollmann, 2009; Hüner et al., 2011).

Insgesamt kann festgestellt werden, dass die Relevanz für eine unternehmensübergreifende Stammdatenqualität aufgrund des identifizierten Bedarfs gegeben ist. Ferner ist festzustellen, dass es derzeit kein Hilfsmittel zur Vereinbarung der erforderlichen Stammdatenqualität zwischen zwei Unternehmen gibt.

Die zentralen Forschungsfragen für das Promotionsvorhaben lauten deshalb:

FF1: Wie muss ein Hilfsmittel gestaltet sein, um weitgehend automatisch die Qualitätsanforderungen von Stammdaten zwischen zwei Unternehmen zu vereinbaren?

FF2: Wie muss ein Hilfsmittel gestaltet sein, um die entsprechenden Stammdaten gemäß dem vereinbarten Qualitätsniveau sicherzustellen?

Der Beitrag leistet keine konkreten Antworten auf diese Fragen, sondern stellt vielmehr den Gegenstand des Promotionsvorhabens dar und zeigt die Vorgehensweise zur Bearbeitung der Forschungsfragen. Der Beitrag ist wie folgt aufgebaut. Im sich der Einleitung anschließenden Kapitel werden Zielsetzung und die zu erwartenden Ergebnisse dargestellt. Kapitel 3 beschreibt die Methodik des Forschungsvorhabens und zeigt in welcher Phase sich das Forschungsvorhaben befindet. Kapitel 4 erläutert exemplarisch unternehmensübergreifende Stammdatenprobleme infolge mangelhafter Datenqualität. Kapitel 3 definiert In Kapitel 5 wird die Funktionsweise des Hilfsmittels für eine unternehmensübergreifende Stammdatenqualität skizziert, bevor der Beitrag mit einem kurzen Fazit und dem Ausblick auf weitere Schritte endet.

2 Zielsetzung

Das Promotionsvorhaben beschäftigt sich im Konkreten mit der Entwicklung und Evaluierung eines Hilfsmittels, mit dem zwei Unternehmen eine Vereinbarung und Prüfung des Qualitätsniveaus für Stammdaten, die zwischen den Unternehmen übertragen werden, weitgehend automatisiert durchführen können.

Das Ziel der Forschungsarbeit ist u. a. die Entwicklung einer Vorgehensweise für eine bilaterale Vereinbarung der Qualitätsanforderungen für Stammdaten zwischen zwei Unternehmen. Zur Prüfung und Sicherstellung der Stammdatenqualität sind Methoden zur Beschreibung, Übermittlung und Überprüfung der Qualitätsanforderungen erforderlich. Für die Evaluierung des Vereinbarungsprozesses und der Bewertung des Qualitätsniveaus für Stammdaten zwischen Unternehmen soll ein Prototyp implementiert werden.

3 Methodik

Die Forschungsarbeit folgt grundsätzlich den Richtlinien der gestaltungsorientierten Wirtschaftsinformatik (Design Science) und enthält die Phasen: Analyse, Entwurf, Evaluation und Diffusion (Hevner et al. 2004, Österle et al. 2010, Peffers et al. 2012).

Erste Untersuchungen mittels Literaturanalyse und Experteninterviews ergaben, dass das Themengebiet zur unternehmensübergreifende Stammdatenqualität noch relativ unbekannt ist und der anvisierte Forschungsbereich noch zu entdecken gilt. Eine prädestinierte Ausgangssituation für einen explorativen Forschungsansatz zur Informationsbeschaffung von ersten Erkenntnissen und Annahmen in diesem Forschungsbereich mittels Fokusgruppeninterviews und Fallstudien.

In einem Fokusgruppeninterview, eine erweiterte Form von Experteninterviews, befragt der Wissenschaftler eine Gruppe von Fachexperten und dokumentiert die Erkenntnisse für eine spätere Auswertung (Morgan 1993; Rosemann und Vessey 2008; Chiarini Tremblay et al. 2010). Ferner dient es zur Gewinnung neuer Einblicke und Ideen und der Strukturierung des Problems in kleine überschaubare Einheiten. Hieraus können dann mögliche Ursachen formuliert werden (Mayerhofer 2007).

Die Fallstudienforschung untersucht in der Regel komplexe, schwer abgrenzbare Phänomene in ihrem praktischen Umfeld (Yin 2002). Sie gilt als eine Form der Aktionsforschung. Wissenschaftler arbeiten gemeinsam mit Fachexperten eines Unternehmens an einer spezifischen Problemlösung (Wilde und Hess 2006, Ofner 2013, Göthlich 2003; Dubé, Paré 2003) und unterstützt die Konstruktion von Artefakten (Eisenhardt 1989).

Das gewählte Forschungsdesign enthält sowohl Experten- und Fokusgruppeninterviews als auch fallstudienbasierte Forschung, in dem die Phänomene erfasst und erste Ansätze von Lösungen in Form von Handlungsempfehlungen gestaltet werden. Folglich basiert der

methodische Ansatz der Forschungsarbeit auf der Aktionsforschung und die Umsetzung erfolgt im engen Austausch mit der Praxis in mehreren iterativen Zyklen der Schritte: Analyse, Aktion, und Evaluation (Wilde und Hess 2006).

Aktuell befindet sich das Forschungsvorhaben in der Analysephase und dort in der Anforderungsanalyse. Für den gegenwärtigen Erkenntnisgewinn wurde eine Literaturanalyse, 33 Experteninterviews, ein Fokusgruppeninterview mit 14 Stammdatenexperten von 12 Unternehmen und 16 Expertenworkshops durchgeführt.

Die Experteninterviews wurden in Form von halbstrukturierten Interviews mit einem Experten durchgeführt. Ein Interviewleitfaden mit 18 Fragen, der iterativ auf Basis der Erfahrung aus den ersten Interviews überarbeitet wurde, diente der Gesprächsstrukturierung sowie der Vollständigkeitsicherung und wurde den Experten ca. eine Woche vor dem jeweiligen Interviewtermin übermittelt. Der vollständige Interviewleitfaden umfasst 18 Fragen. Der Interviewzeitraum der Experteninterview lag zwischen Mitte August bis Anfang November 2015. Es wurden insgesamt ein persönliches (pers.) und 29 telefonische Interview (tel.) durchgeführt. Ein Experte beantwortete die Fragen schriftlich (schriftl.). Die Interviewdauer lag zwischen 42 Minuten und 125 Minuten (durchschnittlich bei etwa 70 Minuten). Die Interviews wurden mit Einverständnis der Experten elektronisch aufgezeichnet und anschließend inhaltsanalytisch ausgewertet (Mayring & Fenzl, 2014). Dabei wurde sowohl eine deduktive als auch induktive Kategorienbildung angewendet, die anschließend mit deskriptiver Statistik nochmals ausgewertet wurde (Hug, Poscheschnik, Lederer, & Perzy, 2015).

Die Expertenworkshops wurden ebenfalls in Form von halbstrukturierten Interviews durchgeführt, jedoch mit 2 bis 5 Experten und beim Unternehmen Vorort. Der Expertenleitfaden enthielt 22 Fragen und zur vollumfänglichen Analyse wurde vom Unternehmen weitere interne Dokumentation, Stammdatenbeschreibung und Excel-Dateien für den Produktstammdatenaustausch zur Verfügung gestellt. Die Workshops wurden im Zeitraum von Juni bis September 2016 durchgeführt. In Summe wurden 16 Expertenworkshops mit 48 Experten der Organisationseinheiten: Stammdatenteam, Einkauf, Vertrieb und IT, durchgeführt. Die Workshops dauerten zwischen 4 bis 6,5 Stunden. Anschließend wurde der ausgefüllte Leitfaden mit weiteren Fragen den Experten zur Verifikation und Ergänzung zur Verfügung gestellt. Die Workshops halfen das in den Experteninterviews bereits identifizierten Problem einer nicht vorhandenen unternehmensübergreifenden Stammdatenqualität im Detail zu beleuchten und dabei die heutigen Prozesse, Mechanismen und Werkzeuge für den Stammdatenaustausch kennenzulernen.

4 Unternehmensübergreifende Stammdatenprobleme

In den Fokusgruppeninterviews und Expertenworkshops konnten gravierende geschäftliche Probleme infolge mangelnder Stammdatenqualität in den Unternehmen identifiziert werden. Eine mangelhafte Stammdatenqualität ist beispielsweise dann gegeben, wenn die Stammdaten

nicht vollständig sind oder fehlerhafte Inhalte enthalten. Im Folgenden werden die häufigsten Probleme mangelhafter Qualität von Produktstammdaten, deren Auswirkung in den Geschäftsprozessen und deren potentiellen wirtschaftlichen Schaden erläutert.

Die **Zolltarifnummer** ist eine essentielle Information bei der Zollabfertigung. Wird diese nicht korrekt in den Produktstammdaten gepflegt, kann die Ware an den Landesgrenzen nicht reibungslos abgefertigt werden. Die Folge ist: der Kunde erhält die Ware nicht zum vereinbarten Liefertermin und bei verderblichen Waren droht der Ausfall. Es entstehen hohe Zusatzkosten, Umsatzeinbußen und Kundenunzufriedenheit.

Abmaße und Gewichte sind wichtige Informationen für logistische Prozesse. Werden diese nicht mit den korrekten Werten in den Produktstammdaten gepflegt, können die Lieferkosten nicht korrekt berechnet werden. Ferner ist eine effiziente Lagerauslastung und ein optimierter Warentransport nicht möglich. In diesen Fällen entstehen ebenfalls hohe Zusatzkosten, Umsatzeinbußen und die Nichtrealisierung neuer Geschäftsmodellen, wie „Same Day Delivery“.

Technische Beschreibungen sind wichtige Informationen für den Warenverkauf. Werden diese nicht detailliert in den Produktstammdaten gepflegt, ist ein Online-Handel eingeschränkt möglich. Die Folge ist eine ergebnislose Onlinesuche und ein Produktvergleich. Die Folge sind Umsatzeinbußen bzw. Kundenunzufriedenheit und -verlust.

Konditionen, Preise oder Warengruppen sind wichtige Informationen im Vertrieb und Einkauf. Werden diese falsch oder unvollständig in den Produktstammdaten erfasst, führt dies zu Problemen in den Geschäftsprozessen. Die Rechnungsprüfung kann nicht durchgeführt werden und eine Lieferantenbewertung im strategischen Einkauf ist nicht möglich. Die Folge sind Zusatzkosten in manueller Nachbearbeitung und voller Lager, da das Verkaufsverhalten der Kunden nicht rechtzeitig bewertet werden konnte.

Insgesamt kann festgestellt werden, dass bei einer mangelhaften Stammdatenqualität erhebliche Umsatzverluste und Imageschäden im Unternehmen entstehen können. Für viele dieser Stammdatenprobleme ist der Fehlerverursacher der (Daten-) Lieferant. Da dieser die angeforderten Stammdaten nicht in ausreichender Qualität dem Datennutzer zur Verfügung stellt bzw. die Anforderung an die Stammdatenqualität durch Datennutzer nicht eineindeutige spezifiziert ist und der Lieferant nach bestem Gewissen die Stammdaten übermittelt.

5 Das Hilfsmittel in Funktion

Im Folgenden wird das Hilfsmittel zur Koordination einer unternehmensübergreifenden Stammdatenqualität in seiner grundsätzlichen Funktionsweise erläutert und die nach jetzigem Kenntnisstand bekannten funktionalen Anforderungen aufgeführt.

Funktionsweise

Das Hilfsmittel dient als Kommunikations- und Prüfinstrument zwischen zwei Unternehmen und unterstützt den bilateralen Vereinbarungsprozess in Hinblick auf das Qualitätsniveau bzw. zur automatisierten Validierung der übermittelnden Stammdaten basierend auf dem vereinbarten Qualitätsvertrag.

Abbildung 1 zeigt schemenhaft das Hilfsmittel mit den zentralen Komponenten: Data Quality Gate und Vereinbarungskomponente. Das Hilfsmittel ist sowohl beim Datennutzer als auch beim Datenlieferant vorhanden und interagiert mit dem jeweiligen unternehmensinternen Stammdaten- bzw. Datenqualitätsmanagement. Zur Übertragung der relevanten Informationen wird zwischen den zwei Hilfsmitteln eine bilaterale Kommunikationsverbindung aufgebaut. Der Grundarchitektur gleicht einem Peer-to-Peer Ansatz. Die Abbildung zeigt ferner in einer starkvereinfachten Art und Weise die zentralen Prozesse zur Vereinbarung bzw. zur Validierung der Stammdatenqualität.

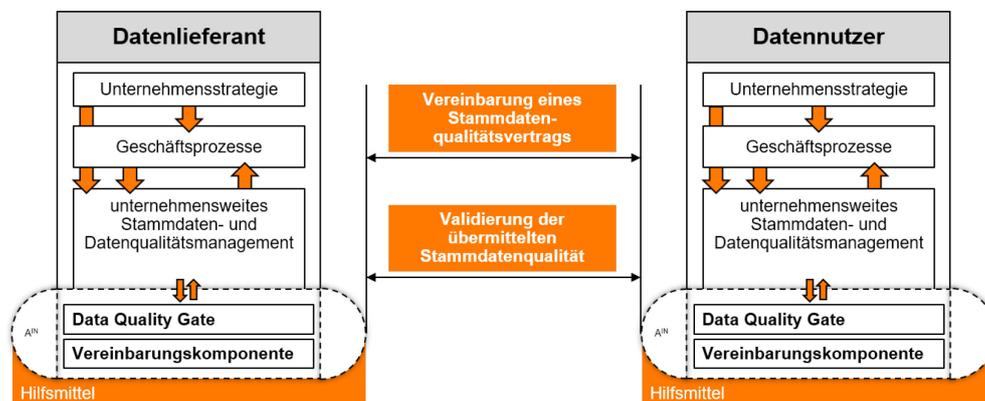


Abbildung 7: Übersicht der Prozesse und Komponenten des Hilfsmittels

Aus der Analyse konnten zur Koordination einer unternehmensübergreifenden Stammdatenqualität zwei aufeinanderfolgende Prozessschritte identifiziert werden:

Prozessschritt 1: *Vereinbarung eines Stammdatenqualitätsvertrags.*

Der Datennutzer kann seine Qualitätsanforderungen an die Stammdaten anwendungsspezifisch formulieren, so dass der Datenlieferant versteht, in welcher Qualität (Syntax, Semantik und Pragmatik) die Stammdaten zu liefern sind.

In einem dafür vorgesehenen Vereinbarungsprozess kann die Qualität weitgehend automatisch zwischen Datennutzer und Datenlieferant ausgehandelt werden. Das Ergebnis wird in einem Stammdatenqualitätsvertrag festgehalten.

Prozessschritt 2: *Validierung der übermittelten Stammdatenqualität.*

Der Datenlieferant kann mittels den Qualitätsanforderungen und Prüfmechanismen die Stammdaten in derselben Art und Weise prüfen, wie der Datennutzer selbst und ggf. Fehlerbeseitigungsmaßnahmen noch vor der Datenübertragung durchführen.

Der Datennutzer erhält beim Stammdatenaustausch erweiterte Informationen über die Qualität der übermittelten Daten (z. B. Prüfbericht vom Datenlieferanten).

Der Datennutzer kann nun auf Basis des Stammdatenqualitätsvertrags nun selbst die übermittelten Stammdaten validieren. Bei Vertragsverletzung könnte die Datenannahme gänzlich verweigern bzw. dedizierte Korrekturmaßnahmen eingeleitet werden. Dies erfolgt vor der Stammdatenintegration in die unternehmensinternen IT-Systeme.

Funktionale Anforderungen

Im Rahmen der Analysephase konnten in Zusammenarbeit mit den Stammdatenexperten einige grundlegende funktionalen Anforderungen an das Hilfsmittel aufgestellt werden. Tabelle 1 zeigt im Überblick die funktionalen Anforderungen an das Hilfsmittel.

Komponente	Funktionale Anforderung
Analyse-Modul	Test und Simulation einer DQ-Prüfung; Bereitstellung eines lieferantenübergreifenden Analysewerkzeug
Cockpit	Visualisierung der Stammdatenqualität anhand vorher definierter Kennzahlen in Form eines Cockpits zur Überwachung
Datenmodell	Bereitstellung von spezifischen Vokabulars; Abbildung von Klassifikationsschemata (eCl@ss)
Prozessmanagement	Artikeldatenaustausch mit Geschäftspartner; Automatisierung und vollautomatisch; Abbildung von Genehmigungsworkflow
Regelwerk	Automatische Generierung von Geschäftsregeln zur Validierung (messen und kontrollieren) der Stammdatenqualität; DQ-Richtlinien; Beschreibungssprachen für Regel; Best Practise-Regeln
Software-Architektur	Form eines modularen Baukasten; Anpassbarkeit; Parametrisierung
Software-Ergonomie	Einfache Handhabung in der Bedienbarkeit; Verständlich;
Software-System	Nutzung bestehender Software-Lösungen wie Excel; Webbasierte Lösung; WebServices
Software-Integration	Rückweg: Aktualisiert sich von alleine / Selbstlernend; System-to-System-Integration (SAP); Integration von Datenbanksysteme; SAP-Schnittstelle; Integrierbar in Warenwirtschaftssystem (z. B. Sage); Einfache Übermittlung der Informationen (Reifegradmodell, Geschäftsregeln, DQ-Index) zum Geschäftspartner; Bereitstellung von maschinenlesbare Information

Tabelle 1: Übersicht der funktionalen Anforderungen an das Hilfsmittel

6 Zusammenfassung und Ausblick

Die bisherigen Analyseergebnisse lassen interessante Sachverhalte erkennen, dass beispielsweise Basistechnologien und Standards (EDI-Verbindungen, Produktidentifikation und Klassifizierungsstandards, Online-Produktkataloge und Datenpools) nur teilweise bis gar nicht den Vereinbarungsprozess einer angemessenen Stammdatenqualität unterstützen. Unternehmen (Datennutzer) einen erheblichen Aufwand haben, die fehlerhaft übermittelten Stammdaten zu identifizieren und zu beheben. Ferner sind Unternehmen meist nicht in der Lage die Anforderungen an die Stammdatenqualität eineindeutig zu spezifizieren bzw. führen keine systematische Qualitätsprüfung der angelieferten Stammdaten durch. Wenn überhaupt, werden manuelle Stichproben durchgeführt.

Ein Hilfsmittel zur bilateralen Vereinbarung und Validierung der Stammdatenqualität wird von den Stammdatenexperten auf dem Weg zu einer weitgehend automatischen Verarbeitung der übermittelten Stammdaten als richtungsweisend eingestuft.

Die nächsten Forschungsaktivitäten ist der Abschluss der Phase „Analyse“, insbesondere der weiteren Vervollständigung der Anforderungsanalyse und anschließenden Priorisierung der funktionalen Anforderungen nach der Methode Quality Function Deployment (QFD).

Anschließend folgt die Phase 2 „Entwurf“ mit der Entwicklung eines bilateralen Vereinbarungsprozesses und der Konstruktion eines Prototyps gemäß der Anforderungsanalyse. Zur iterativen Gestaltung und Anpassung der Artefakte sind folgende Techniken angedacht: Methoden-Engineering, Referenzmodellierung und Software-Engineering.

Literaturverzeichnis

- Bakos, J. Y. (1991). Information links and electronic marketplaces: the role of interorganizational information systems in vertical markets. *Journal of Management Information Systems*, 8(2), 31–52.
- Bloching, B., Leutiger, P., Oltmanns, T., Rossbach, C., Schlick, T., Remane, G., . . . Shafranyuk, O. (2015). Die digitale Transformation der Industrie: Detailbetrachtungenn von Rolannd Berger Strategy Consultants im Auftrag des Bundesverbands der Deutschen Industrie e.V.(BDI). München.
- Dalmolen, S., Moonen, H., & van Hillegersberg, J. (2015). Industry-wide Inter-organizational Systems and Data Quality: Exploratory findings of the use of GS1 standards in the Dutch retail market. In *Twenty-first Americas Conference on Information Systems*.
- Davenport, T. H., & Short, J. E. (1990). The New Industrial Engineering: Information Technology and Business Process Redesign. *Sloan Management Review*. (4), 11–27.
- de Corbière, F. (2007). Interorganizational Information Systems and Data Quality Improvement: The Case of Product Information in the French Large Retail Industry. In M. A. Robbert, R. O'Hare, M. L. Markus, & D. K. Barbara (Eds.), *Proceedings of the 12th International Conference on Information Quality*. Mary Ann Robbert, Robert O'Hare, M. Lynne Markus, Barbara D. Klein.:

- GS1 Germany. (2016). Global Data Synchronisation Network (GDSN). Retrieved from <http://www.gs1.org/gdsn>
- Haug, A., & Stentoft Arlbjørn, J. (2011). Barriers to master data quality. *Journal of Enterprise Information Management*, 24(3), 288–303.
- Helmis, S., & Hollmann, R. (2009). *Webbasierte Datenintegration: Ansätze zur Messung und Sicherung der Informationsqualität in heterogenen Datenbeständen unter Verwendung eines vollständig webbasierten Werkzeuges* (1. Aufl.). *Ausgezeichnete Arbeiten zur Informationsqualität*. Wiesbaden: Vieweg+Teubner Verlag / GWV Fachverlage GmbH Wiesbaden.
- Hug, T., Poscheschnik, G., Lederer, B., & Perzy, A. (2015). *Empirisch forschen: Die Planung und Umsetzung von Projekten im Studium* (2., überarb. Aufl.). *Studieren, aber richtig: Vol. 3357*. Konstanz: UVK-Verl.-Ges.
- Hüner, K. M., Schierning, A., Otto, B., & Österle, H. (2011). Product data quality in supply chains: The case of Beiersdorf. *Electronic Markets*, 21(2), 141–154.
- Legner, C., & Schemm, J. (2008). Toward the Inter-organizational Product Information Supply Chain - Evidence from the Retail and Consumer Goods Industries. *Journal of the Association for Information Systems*, 9(3/4), 119–150.
- Legner, C., & Wende, K. (2007). The Challenges of Inter-Organizational Business Process Design - A Research Agenda. In H. Österle & J. Schelp (Eds.), *Proceedings of the 15th European Conference on Information Systems (ECIS 2007)* (pp. 1643–1654). St. Gallen.
- Loshin, D. (2008). *Master Data Management*. Amsterdam: Morgan Kaufmann.
- Madlberger, M. (2011). Can data quality help overcome the penguin effect? The case of item master data pools. In *ECIS 2011 Proceedings*.
- Mayring, P., & Fenzl, T. (2014). *Qualitative Inhaltsanalyse*. In N. Baur & J. Blasius (Eds.), *SpringerLink : Bücher. Handbuch Methoden der empirischen Sozialforschung* (pp. 543–556). Wiesbaden: Springer
- Ofner, M., Straub, K., Otto, B., & Oesterle, H. (2013). Management of the master data lifecycle: a framework for analysis. *Journal of Enterprise Information Management*, 26(4), 472–491.
- Ofner, M. H., Otto, B., & Österle, H. (2012). Integrating a data quality perspective into business process management. *Business Process Management Journal*, 18(6), 1036–1067.
- Otto, B., Abraham, R., & Schlosser, S. (2014). Toward a Taxonomy of the Data Resource in the Networked Industry. In *7th International Scientific Symposium on Logisitics* (pp. 382–421).
- Otto, B., & Österle, H. (2016). *Corporate Data Quality*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Schemm, J. W. (2009). *Zwischenbetriebliches Stammdatenmanagement: Lösungen für die Datensynchronisation zwischen Handel und Konsumgüterindustrie*. SpringerLink : Bücher. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Wand, Y., & Wang, R. Y. (1996). Anchoring data quality dimensions in ontological foundations. *Communications of the ACM*, 39(11), 86–95.
- Wang, R. Y. (1998). A Product Perspective on Total Tata Quality Management. *Communications of the ACM*, 41(2), 58–65.
- Wang, R. Y., & Strong, D. M. (1996). Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems*, 12(4), 5–33.

Kontaktinformationen

Dipl.-Inf. Thomas Schäffer
Hochschule Heilbronn
Fakultät Wirtschaft und Verkehr
Studiengang Wirtschaftsinformatik

Telefon: +49 (0)7131 - 504 6699

E-Mail: thomas.schaeffer@hs-heilbronn.de

Web: www.hs-heilbronn.de

Technische Universität Chemnitz
Fakultät für Wirtschaftswissenschaften
Professur für Wirtschaftsinformatik – Geschäftsprozess- und
Informationsmanagement
Frau Prof. Dr. Barbara Dinter
Thüringer Weg 7
D-09126 Chemnitz
Telefon: +49 (0) 371 531 39228
Fax: +49 (0) 371 531 26529
E-Mail: barbara.dinter@wirtschaft.tu-chemnitz.de
Homepage: <https://www.tu-chemnitz.de/wirtschaft/wi1/>



TECHNISCHE UNIVERSITÄT
CHEMNITZ

TU Chemnitz
09107 Chemnitz
www.tu-chemnitz.de