The Goat Criteria—A Structured Assessment Approach for Reference Models

Christian Janiesch and Axel Winkelmann

1 Introduction

Reference modelling has been an ansatz in Information Systems research since the inception of the discipline (Becker, Niehaves, & Knackstedt, 2004; Fettke & Loos, 2007; Schwegmann, 1999; Thomas, 2006; vom Brocke, 2003). Reference models are generic models for a class of applications with certain properties: They can be the basis for specific models for a particular application and they can be references for comparisons, for example to benchmark one's structure or operations. Further, they can be used to educate and improve communication. They range from abstract frameworks and blueprints to ontologies.

Applying reference models goes beyond changing the exterior functional texture of companies, but they allow for structuring their business architecture as well as their process organizations. Hence, the leitmotif "*structure, organize, texture*" (in German: strukturieren, strukturieren, strukturieren).

The gestalt of reference models is diverse, the most common and extensive models focus on processes and data. In the last 50 years, the scientific community has provided a plethora of reference models (e.g., the "integriertes Gesamtmodell der betrieblichen Datenverarbeitung") (Grochla & Szypersk, 1971). Later works, for example those provided by Scheer or Kurbel (cf. e.g. Scheer, 1997), added valuable knowledge to the IS community.

In science and in particular in Münster, milestones of reference modelling research have been made culminating in eclectic reference models such as the Retail-H by Becker and Schütte (2004)—ranging back to works in 1993–1996—has provided

C. Janiesch (✉) · A. Winkelmann
Julius-Maximilians-Universität Würzburg, Würzburg, Germany
e-mail: christian.janiesch@uni-wuerzburg.de

A. Winkelmann
e-mail: axel.winkelmann@uni-wuerzburg.de

K. Bergener et al. (eds.), *The Art of Structuring*,
https://doi.org/10.1007/978-3-030-06234-7_7

significant value not only to the scientific but also to the IS community by capturing the zeitgeist of contemporary retail business processes and data.

In practice, especially SAP AG provided a stable reference model for its SAP R/3 that served the IT management forces in enterprises for two decades. In 2019, the community still talks about attributes such as lifnr (Lieferantennummer/ supplier number) or tables such as lfA1 (Lieferantenebene A1/ supplier level A1) in the respective systems. While the abbreviations are obvious for German enterprises, there is still confusion e.g. among foreign enterprises using SAP software. Recently, especially Microsoft (2017) introduced a new reference (data) model for enterprises, the Common Data Model (CDM). Microsoft has heavily invested into its Cloud strategy and its reference model is a major part of their ERP movements.

As mentioned above, a primary use case for a reference model is structuring enterprise data and the organizing corporate processes by providing a customizable blueprint or master model to initiate organizational change (vom Brocke, 2007). Yet, one question remains difficult to answer: Which reference model performs better in the long run. Here, companies should not focus on short-term gains but rather consider long-term achievements, which only start to pay off in the mid-term. This leads to the problem, that reference models need to be evaluated using strategic criteria rather than tactical or operational metrics (Fettke & Loos, 2003; Schütte, 2013).

In this festschrift, we introduce a new qualitative measurement scale to evaluate long-term perfection of reference models called the GOAT criteria. The term GOAT (also G.O.A.T.) has been popularized in another discipline by LL Cool J (2000).[1] Our GOAT criteria assess whether a reference model is *generic*, *objective*, *agnostic*, and *transferable*, which are important factors for enduring reference model success. They predicate whether a reference model indeed focuses on a class of applications rather than specialized circumstances, whether it abstains from unnecessary subjectivism, if it is agnostic of technology and implementation choices, and—last but not least—whether it is still practical to be applied in the real-world.

We use these above criteria to evaluate the following reference models:

(a) the Retail-H reference model of Becker and Schütte (2004), which is an established and proven conceptual specification of retail functions, processes, and data.
(b) the Microsoft Common Data Model of Microsoft (2017), which is a library for enterprise data elements as a representative of modern reference models.

Finally, we discuss our findings and summarize the results.

[1]The term GOAT is used in professional sports as well. It is attached to players such as Roger Federer, Wayne Gretzky, Tom Brady. We propose to attach the term GOAT to Jörg Becker for his achievement in reference modelling.

2 GOAT Criteria

2.1 Overview

Reference models can incorporate the constituent parts of any consistent idea, from system modules to business functions. Their common denominator is that they consist of a complete set of items for the reference domain. This frame of reference can then be employed design systems and data structures, initiate business process change management, or to communicate notions unambiguously among members of the same community.

Since reference models vary greatly in their scope and focus, it is impossible to develop detailed and still fair universal measurement catalogues across domains and applications. Hence, in the following we focus on providing high-level criteria, which can be used to assess the perfection of the reference model's ability to structure, organize, and texture enterprises.

Figure 1 provides an overview of the GOAT criteria, which are in detail: generic, objective, agnostic, and transferable. Each of the four dimensions can consist of assessment specific sub-dimensions, which we suggest to structure in a hierarchical fashion to improve clarity. We assume that models can be optimized for at most three dimensions while the fourth dimension introduces a trade-off.

2.2 Generic

A reference model has to be generic rather than specific. It has to provide information about circumstances of interest on an abstract level. A reference model has to describe the kinds or types of entities that may or may not exist in such a circumstance. A reference mode is not about specific entities and their attributes and relations that occur in a specific circumstance. It is not an ersatz for a company- or implementation-specific model.

For example: A reference model may specify that a sports game pitch needs goals and explain the concept of a goal by stating that the formative object of the sport (e.g. a ball) needs to travel in-between the posts to receive an advantage over the competitor (i.e. to score a goal). However, it does not specify exact dimensions or whether the goal needs to be equipped with a net as goals differ from sport to sport (e.g. compare goals in water polo to elephant polo).

That is, a reference model should describe types of entities and their relationships rather than include specific and detailed constraints. The more details a reference models exhibits, the more likely it is that it needs to be customized and that it cannot be applied consistently in its entirety. Further, for any reference model not only to structure data but also to organize a business, precedence relations between roles, tasks and functions need to be made explicate in a process description.

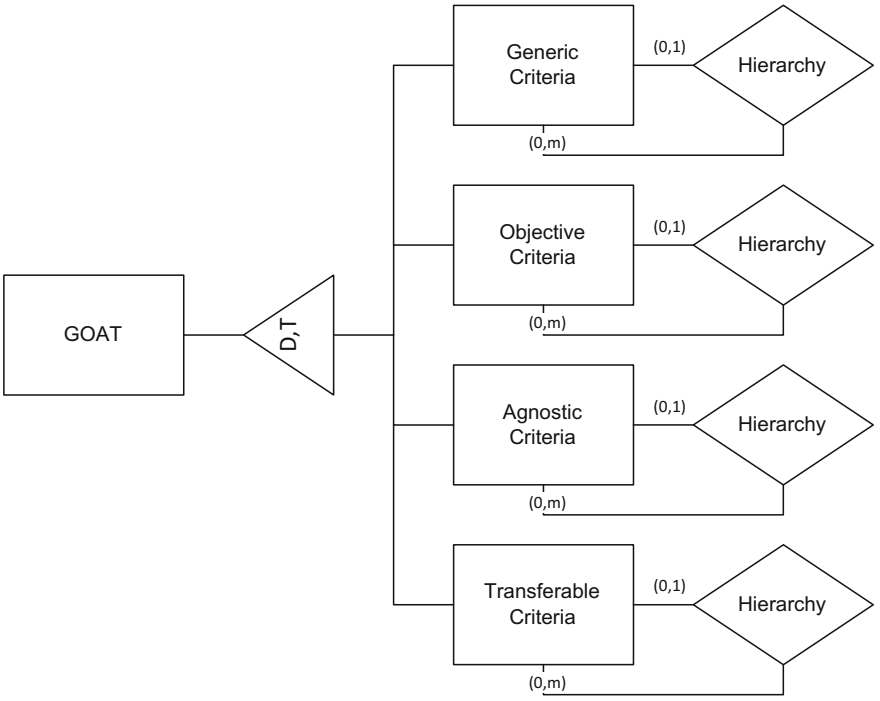

Fig. 1 Conceptual overview of the GOAT criteria

Concerning *composition*, the dimension *generic* of the GOAT criteria can be evaluated by assessing whether the reference model contains a general-purpose data model with entities and relationships as well as a high-level process description with tasks, roles and precedence relations.

Concerning *content*, the dimension *generic* of the GOAT criteria can be evaluated by assessing whether the reference model contains types of entities and tasks as well as their relations rather than specific instantiations thereof. If distinctions for cases have to be made within the model, generic meta categories should be employed to refrain from using situational denominations, which need to be replaced for use in other circumstances.

2.3 Objective

Objectivity informs essential processes in science and how evidence and truth is unearthed. Its central conception is to eliminate personal biases and to channel reproducible results. Objective reference models can be associated with a truth value and are in line with the objective truth as for example discussed in the Habermas Theory

of Communicative Action (Habermas, 1984). For a discourse leading to a cre-dible result, i.e. the use of a reference model, the theory postulates, that apart other things the content of the message, i.e. the reference model, has to represent the circumstances it is thought to describe truthfully. An explicated research method associated to the reference model can provide this clarity.

The topic of objective truth is linked to the quality of information in the models for a realistic and error-free reproduction of the circumstances it is thought to describe. Quality can be based on the ability to satisfy declared or implied needs, based on its totality of characteristics (DIN, 2015). Wang and Strong (1996) provide a framework of dimensions for information quality with criteria to measure. They distinguish the dimensions of intrinsic, contextual, representational, and accessibility. Further criteria have been introduced by Schütte (2013) when extending the Guidelines of Modeling (Schütte & Rotthowe, 1998). For reference models it is particularly important to follow an established method, such as established general-purpose modeling languages as ERM (Chen, 1976) and BPMN (Object Management Group Inc., 2013).

Concerning *truth*, the dimension *objective* of the GOAT criteria can be evaluated by assessing whether the reference model contains an explicated research procedure to make the development process reproducible and whether it makes raw data available.

Concerning *quality*, the dimension *objective* of the GOAT criteria can be evaluated by assessing whether the reference model adheres to open quality standards such as established modelling languages, modelling guidelines, and/ or information quality criteria and integrates them in a homogenous and holistic fashion.

2.4 Agnostic

A reference model's longevity is limited if it necessitates certain technologies or platforms. A reference model's intent is to promote understanding and solutions a type of circumstances, not specific solutions specific situations. Hence, it should be agnostic of any technological or political constraint, which could—in the long run—entail a predicament. While reference models exist on various levels of abstraction, they should not include concrete implementations as otherwise the distinction towards technology specifications, which describe concrete technological solutions, is unsustainable.

Furthermore, the reference model should refrain from overly aligning with current and not well-understood trends. By seeking closeness to hyperbolic topics, a reference model may excel at appearing to be superior to previous revisions or older models. Yet, the shelf life of the reference model in question will most likely not benefit from this.

Concerning *technology*, the dimension *agnostic* of the GOAT criteria can be evaluated by assessing whether the reference model necessitates concrete technologies rather than generalizable IT-based functionality.

Concerning *hyperbole*, the dimension *agnostic* of the GOAT criteria can be evaluated by assessing whether the reference model contains an excessive number of current catchwords, which do not translate into distinguishable benefits.

2.5 Transferable

To be useful for reuse and implementation, a reference model need to be contextualized to specific situations or contexts when applied in practise. This dimension, thus, introduces a trade-off. While models need to be generic in their nature, objectively formulated, and agnostic in relation to their implementation to ensure enduring relevance, reference models need to be transferable into practise to be useful to the scientific community and—more importantly—to practice. There exists a body of research on reference model contextualization describing mechanisms that can assist this process (Delfmann, 2006).

Further, reference models should include a clear statement of the circumstances that it applies to and the problems it solves. For a reference model to be a relevant contribution to the community, it needs to solve a practical problem such as the design of a generic IT system for business functionality. Further, this problem needs to be non-trivial and the solution non-obvious, otherwise routine design should be preferred (Gregor & Hevner, 2013).

Concerning *adaptability*, the dimension *transferable* of the GOAT criteria can be evaluated by assessing how much adaptation in terms of customization and instantiation the reference model requires to be used in practise. Providing configuration mechanisms can alleviate the contextualization process.

Concerning *practicability*, the dimension *transferable* of the GOAT criteria can be evaluated by assessing whether the reference model solves a practical problem in a domain where solutions are non-obvious.

3 Reference Models for Enterprise Data

3.1 Retail-H

At the beginning of the 1990s, Jörg Becker and his research team conducted several research and consulting retail projects in the context of business information systems. With their aim to structure, organize, texture, they soon realized that a retailing company must perform three principal tasks, namely procurement, storage, and distribution of goods. This was an initial step of building the reference model Retail-H (German: “Handels-H”). All tasks concerned with the supplier are addressed within the “left leg”, all tasks associated with the customers within the “right leg” and

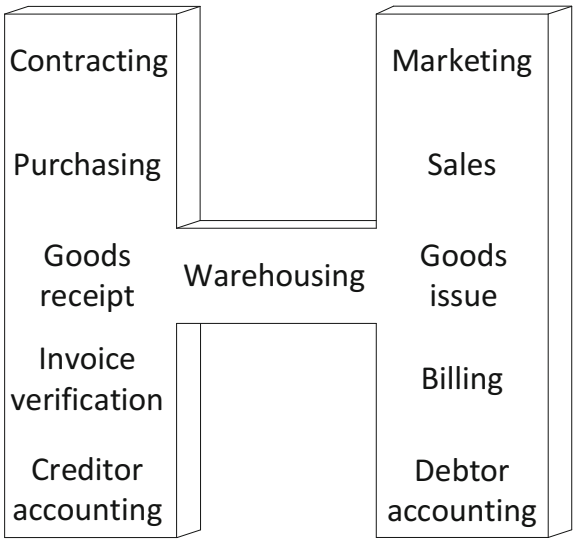

Fig. 2 Retailing H structure

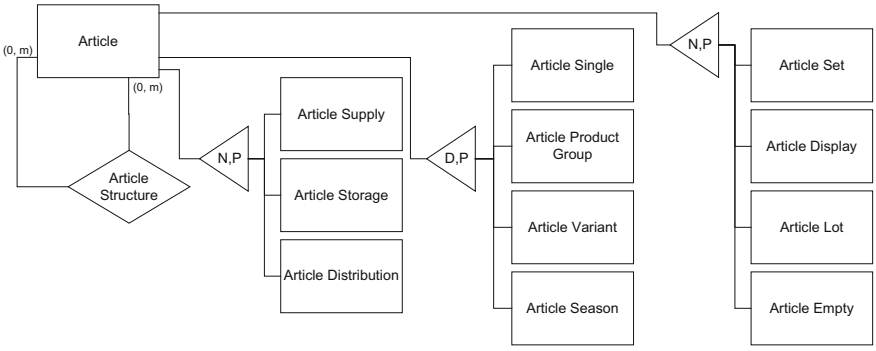

Fig. 3 Retail-H reference data model (cf. Becker & Schütte, 2004)

logistical functions with goods receipt, warehousing and goods issue are arranged horizontally. Hence, the structure of the retailing functions forms an H (cf. Fig. 2).

These tasks of procurement, distribution and logistics occur not only in retailing companies but also in industrial companies. However, retailing has some unique features that affect the form of the tasks and consequently the business information systems. Therefore, Becker and Schütte (2004) provided additional layers underneath the H structure in order to specify functions, processes and data needed for business information systems in retail. Next to functional decomposition diagrams and event-driven process chains (EPC) for reference process models, they provided various retail constructs. These ERM data models provide valuable insight into the necessary data specifications in retail (cf. Fig. 3).

3.2 Microsoft Common Data Model

Microsoft has changed its business model intensively during the last couple of years (Nadella, Shaw, & Nichols, 2017). It has become the world's #1 cloud provider and has spent much effort not only to provide its Office solutions within its Azure cloud but also its business software world. Lately, Microsoft has heavily invested into reconfiguring and reshaping its now cloud based ERP systems Microsoft Dynamics Business Edition and Enterprise Edition (both formerly known as NAV/Navision and AX/Axapta). A rigorous service orientation and modularization (in combination with Apps, Azure services, CRM services, LinkedIn, etc.) will provide more flexibility to customers in the future.

To enable service providers to understand and follow Microsoft's new approach, the company has defined a reference data model (cf. Fig. 4). The so-called Common Data Model (CDM) is an open-source definition of standard entities that represent commonly used concepts and activities across a variety of business and application domains within the Microsoft world. Their public CDM GitHub repository will be continuously enhanced with core entities spanning the entire business process landscape, additional vertical industry data models, and cross-spanning sources such as surveys, search engines, and product telemetry.[2] CDM offers well-defined, modular, and extensible business entities such as Account, Business Unit, Case, Contact, Lead, Opportunity, and Product, as well as interactions and relationships between vendors, workers, and customers, such as activities and service level agreements.

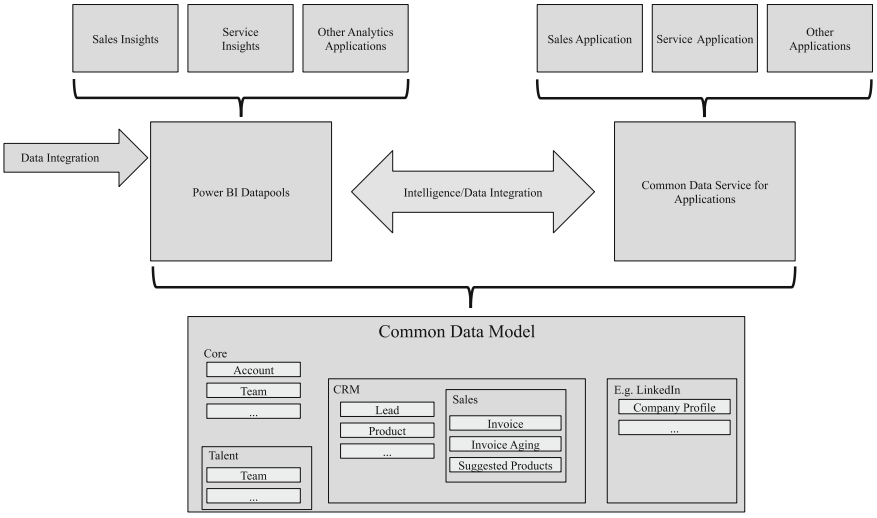

Fig. 4 Microsoft's common data reference model (cf. Microsoft Corporation, 2018)

[2] See https://github.com/Microsoft/CDM/.

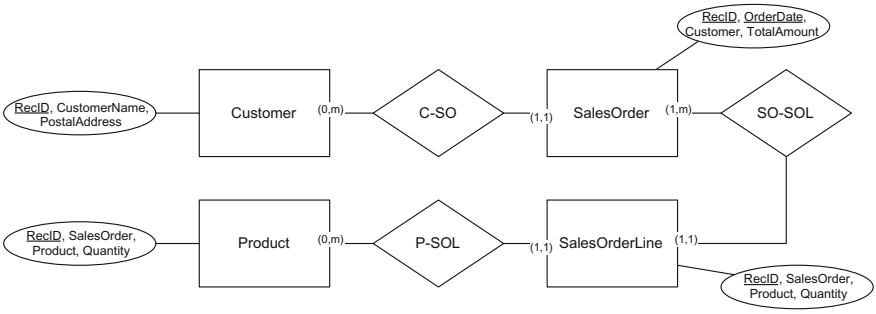

Fig. 5 CDM business entity "sales order" (cf. Microsoft Corporation, 2017)

CDM promises to overcome the challenges of data management by unifying data in a known form with structural and semantic consistency across applications and deployments. It will help integrate and disambiguate data collected from among others business processes, digital interactions, product telemetry, and people interactions. For each entity group, there is an ERM that describes the entities and their relationships with each other. The following ERM shows a—simplified—diagram for a sales order (Fig. 5).

4 Discussion and Conclusion

In the following, we provide a blueprint for a first assessment of the Retail-H and the Microsoft CDM as a representative of contemporary company-originated reference models. We have made a brief survey of the entirety of the two reference models and highlight our initial findings. We have arranged all criteria in Table 1.

Both models make extensive use of general-purpose modelling languages. From a scientific standpoint, the Retail-H provides a much more consistent and generic structure and content. Similarly, the Retail-H's development and description is based on a scientific process, while the CDM's content is not independently verifiable. Both provide high-quality conceptual models though. Further, the Retail-H is completely technology-agnostic while the CDM favours certain technologies. Both model refrain from overstatements. In terms of the being transferable, the CDM clearly has some advantages but its objective and agnostic limitations constrain its genericity toward the Microsoft-affiliated world. It is most likely kaput in other contexts.

As introduced early, it is impossible for a reference model to excel in all four dimensions. Hence, the GOAT criteria always have a qualitative component, which needs to value to different dimension's degree of fulfilment and their importance.

In this case—the case of longevity and versatility–, it is rather obvious that the first three dimensions should be valued at a higher ratio. Here, the Retail-H outperforms the CDM clearly and it lets one assume that the Retail-H will still be talked about

Table 1 Comparison of reference models using the GOAT criteria

GOAT	Retail-H	CDM
Generic	Structured along process, data, and function models; content is clearly generic and on type level and linked through complex relations	Data model only; content is mostly on type level but not fully normalized (e.g. SocialNetwork01, SocialNetwork02) and contemporary rather than generic (e.g. LinkedInIdentity, TwitterIdentity)
Objective	Stands on the shoulders of giants, many references; high-quality data models based on ERM, processes in EPC, further function and organization model of ARIS	No published research method, thus truth value unknown; high-quality data models based on ERM
Agnostic	Completely technology agnostic; no hyperboles	Links to Microsoft and third-party technologies (LinkedIn, Twitter), strong link to Microsoft Power Apps; no hyperboles
Transferable	High instantiation efforts necessary, but generic adaptation mechanisms are available for the included modelling languages; its practical use has been proven times over in the last 20 years	Low to medium instantiation efforts necessary, but only within the Microsoft Power Apps domain, concrete adaptability is thus unclear; but its practical use can be assumed

for the next 60 years to come while CDM's future is unclear and depending on the technological choices made implicitly and explicitly in the model.

Thus, we can conclude that the Retail-H represents a strong reference model with a high degree of perfection if not the GOAT of reference models. This analysis, however, is subject to future research.

References

Becker, J., Niehaves, B., & Knackstedt, R. (2004). Bezugsrahmen zur epistemologischen Positionierung der Referenzmodellierung. In J. Becker & P. Delfmann (Eds.), *Referenzmodellierung: Grundlagen, Techniken und domänenbezogene Anwednung* (pp. 1–17). Heidelberg: Physica.

Becker, J., & Schütte, R. (2004). *Handelsinformationssysteme: Domänenorientierte Einführung in die Wirtschaftsinformatik* (2nd ed.). Frankfurt am Main: Redline Wirtschaft.

Chen, P. P.-S. (1976). The entity-relationship model—toward a unified view of data. *ACM Transactions on Database Systems, 1*(1), 9–36.

Delfmann, P. (2006). *Adaptive Referenzmodellierung: Methodische Konzepte zur Konstruktion und Anwendung wieder verwendungsorientierter Informationsmodelle*. Berlin: Logos Verlag.

DIN. (2015). *Quality management systems - requirements* (ISO 9001:2015).
Fettke, P., & Loos, P. (2003). Classification of reference models: A methodology and its application. *Information Systems and e-Business Management, 1*(1), 35–53.
Fettke, P., & Loos, P. (2007). Perspectives on Reference Modeling. In P. Fettke & P. Loos (Eds.), *Reference modeling for business systems analysis* (pp. 1–20). Hershey, PA/London: Idea Group.
Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. *MIS Quarterly, 37*(2), 337–355.
Grochla, E., & Szypersk, N. (Eds.). (1971). *Management-Informationssysteme: Eine Herausforderung an Forschung und Entwicklung*. Wiesbaden: Springer Fachmedien Wiesbaden.
Habermas, J. (1984). *Theory of communicative action: Reason and the rationalization of society* (p. 1). Boston: Beacon Press.
LL Cool, J. (2000). *G.O.A.T. (Greatest of All Time)*. Santa Monica, CA: Def Jam, Universal.
Microsoft Corporation. (2017). Common data model and data integration. Retrieved from https://docs.microsoft.com/en-us/business-applications-release-notes/april18/data-integration/common-data-model-data-integration.
Microsoft Corporation. (2018). *Common data model and data integration*. Retrieved from https://docs.microsoft.com/en-us/business-applicationsrelease-notes/april18/data-integration/common-data-model-data-integration.
Nadella, S., Shaw, G., & Nichols, J. T. (2017). *Hit refresh*. Kulmbach: Börsenmedien AG.
Object Management Group Inc. (2013). *Business process model and notation (BPMN) Version 2.0.2*.
Scheer, A.-W. (1997). *Wirtschaftsinformatik: Referenzmodelle für industrielle Geschäftsprozesse* (Aufl. 7). Berlin: Springer Verlag.
Schütte, R. (2013). *Grundsätze ordnungsmäßiger Referenzmodellierung: Konstruktion konfigurations- und anpassungsorientierter Modelle*. Springer Fachmedien Wiesbaden.
Schütte, R., & Rotthowe, T. (1998). *The guidelines of modeling–an approach to enhance the quality in information models*. Paper presented at the 17th International Conference on Conceptual Modeling (ER) (pp. 240–254). Singapore (SG).
Schwegmann, A. (1999). *Objektorientierte Referenzmodellierung: Theoretische Grundlagen und praktische Anwendung*. Wiesbaden: Deutscher Universitäts-Verlag.
Thomas, O. (2006). Das Referenzmodellverständnis in der Wirtschaftsinformatik: Historie, Literaturanalyse und Begriffsexplikation. In P. Loos (Ed.), *Veröffentlichungen des Instituts für Wirtschaftsinformatik* (Vol. 187). Saarbrücken: Universität des Saarlandes.
vom Brocke, J. (2003). *Referenzmodellierung: Gestaltung und Verteilung von Konstruktionsprozessen*. Berlin: Logos Verlag.
vom Brocke, J. (2007). Design principles for reference modelling: reusing information models by means of aggregation, specialisation, instantiation, and analogy. In P. Fettke & P. Loos (Eds.), *Reference modeling for business systems analysis* (pp. 47–75). Hershey, PA/London: Idea Group.
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.

Christian Janiesch is Assistant Professor at the Julius-Maximillian's-University of Wuerzburg, Germany. He received his Ph.D., M.Sc., and B.Sc. in Information Systems from the University of Münster. Christian worked full-time at the European Research Center for Information Systems (ERCIS) in Münster, at the SAP Research Center Brisbane at SAP Australia Pty Ltd, and at the Karlsruhe Institute of Technology (KIT). He is a member of the flying faculty of the Vietnamese-German University in Ho Chi Minh City, Vietnam. His research is at the intersection of Business Process Management and Business Analytics in particular in the area of Cyber-physical

Systems. He is on the Department Editorial Board for BISE. He has authored over 100 scholarly publications. His work has appeared in BISE, FGCS, CAIS, ISeB, BPMJ, as well as various major international conferences including ICIS, ECIS, BPM, and HICSS and has been registered as U.S. patents.

Axel Winkelmann is a full professor of Information Systems at the University of Wuerzburg. His research interests focus on the various aspects of enterprise systems, especially Enterprise Resource Planning (ERP), including Digital Innovation and Transformation as well as Process Management. With his team of currently 13 Ph.D. students he conducts numerous research projects in the context of Process Management, ERP, Artificial Intelligence and Blockchain mechanisms. Many of his articles and books have been written in co-authorship with Joerg Becker. From 2004–2009, he became the first Managing Director of ERCIS, before he moved to the University of Koblenz and the University of Wuerzburg subsequently.