Inductive Reference Model Development: Recent Results and Current Challenges

Jana-Rebecca Rehse, Philip Hake, Peter Fettke, and Peter Loos[1]

Abstract: Reference modeling offers attractive benefits for both research and practice. The inductive strategy for reference model development derives reference models by generalizing individual enterprise models. It has recently gained attention in research, however, its practical application still faces numerous challenges. The objective of the article at hand is to identify recent contributions to the field of inductive reference model development and use them to analyze the current challenges that impede the application of their results in practice. We identify a total of 18 contributions, either scientific articles describing inductive methods for reference model development, or practical reports describing the concrete development of a reference model for a certain domain. They are all analyzed by means of a six-stage-framework for reference model development. For each stage, we derive specific challenges and point out acknowledgments and potential solutions.

Keywords: Inductive Reference Model Development, Reference Modeling, Grand Challenges

1 Introduction

Reference modeling offers attractive benefits for both research and practice [FL07]. Following the epistemologically established differentiation between rationalism and empiricism as two fundamental ways of cognition, reference modeling differentiates a deductive and an inductive strategy for reference model development [BS97]. Model development according to the deductive strategy employs generally accepted theories and principles, while the inductive strategy is based on the generalization of individual enterprise models. It focuses on the commonalities of the individual models and abstracts from specific features. Hence, deductive development proceeds from the general towards the specific ("top-down"), whereas inductive development evolves from the specific into the general ("bottom-up").

Although several concepts, methods, and tools for the support of reference modeling exist by now, practical development of reference models still faces a variety of challenges. The open questions range from project preparation, over individual steps of pre-processing and derivation, to maintenance and continuous improvement of a reference model. Hence, the development process of a reference model today is often barely structured, nontransparent, and only marginally justified and thus hardly communicable in terms of course and characteristics. In other words, the ideal of a generally repeatable, engineering-style approach is usually not yet reached. With that, a tool support is also only possible in a limited way.

[1] Institute of Information Systems at the DFKI, Campus D3.2, 66123 Saarbruecken, Jana-Rebecca.Rehse@iwi.dfki.de, Philip.Hake@iwi.dfki.de, Peter.Fettke@iwi.dfki.de, Peter.Loos@iwi.dfki.de

The article at hand intends to pave the way for a structured, methodical, repeatable, and thus justifiable practical application of inductive reference modeling techniques. Therefore, we analyze recent contributions to the field of inductive reference model development and point out the current challenges that impede the application of their results in practice. We consider both scientific articles focusing on the underlying methodology and reports describing the concrete inductive development of a domain-specific reference model in practice, intending to identify as many practical challenges as possible.

The challenges as presented in the following are comparable to the currently discussed Grand Challenges of Mertens and Barbian [MB15]. While they discuss Grand Challenges of Business and Information Systems Engineering (BISE) or “Wirtschaftsinformatik” in general, the contribution at hand is limited to the specific challenges of inductive reference modeling. Although width and relevance of the challenges discussed here are much narrower in the first place, we also intend to fruitfully stimulate future research. Like Mertens and Barbian, we do not assume our list to be complete, but rather a first attempt towards a more structural examination of inductive reference model development.

The article is structured as follows. After the introductory section, important foundations of inductive reference modeling are described in Section 2. Section 3 outlines the identification of challenges by means of a literature review. The identified challenges are grouped and described in section 4. Section 5 critically assessed our findings, before concluding the article with an outlook on further research.

2 Foundations of Inductive Reference Model Development

The research agenda for reference modeling can be outlined by two main questions:

1. Questions regarding Reference Models: How to design a reference model for a certain application domain? A reference model can capture the current organizational context (descriptive interpretation) or make a suggestion for an innovative organizational context (prescriptive interpretation).
2. Questions regarding Reference Modeling Methods: How to develop and apply reference models? Methodological questions can examine characteristics and performance features of known methods, refine existing methods, or invent and evaluate completely new methods.

Both research questions can be viewed regarding the inductive development strategy. Answers concerning the current state of research are exemplified by the following aspects:

1. Reference Models: Several works report on the inductive development of a reference model ([AFF11, KPR08, DM05], q.v. the attribute “construction method” of the reference models listed in the catalog at *http://rmk.iwi.uni-sb.de*).

2. Reference Modeling Methods: Known procedure models for reference modeling indicate that existing individual enterprise models as well as further sources of knowledge should be identified and considered within the context of reference modeling [Be02, Th06]. In addition, there are several contributions that suggest methodological approaches for different aspects of inductive reference model development [Ar13, GvJV08, LRW09, RFL13].

Although the research agenda for inductive reference modeling has made a certain progress, there exists a variety of challenges, which have not yet been systematically addressed. In order to provide an extensive and systematic assessment of challenges, we follow the method for inductive reference model development, as defined by Fettke [Fe14]. It consists of six consecutive stages, each identifying and describing a central task necessary for the practical development of a reference model. The stages are shown in Fig. 1 and shortly outlined in the following.

Fig. 1: Stages of Inductive Reference Modeling [Fe14]

1. Preparation of Reference Model Development: Before a reference model can be developed, several preparatory decisions and assumptions have to be made regarding the context of the reference model development. Among other things, this stage includes identifying stakeholders' requirements and defining modeling conventions.

2. Collection of Individual Models: The inductive development strategy is based on individual enterprise models as obligatory input data. These models are collected in this stage. Depending on the size of the reference model domain, the collection may be split into several sub-stages, identifying classes of models and representative organizations before collecting the actual models.

3. Preprocessing of Individual Models: After the individual models have been collected, they have to be aligned and harmonized before they can be used for the derivation of a reference model. This stage includes checking and establishing necessary modeling conventions as well as identifying correspondences.

4. Acquisition of the Reference Model: This stage describes inductive reference model development in a narrower sense. The preprocessed individual models are used to derive a reference model, using an arbitrary similarity measure and construction approach. Depending on the number of input models, it may make sense to split them into several model clusters and develop an individual reference model for each cluster.

5. Postprocessing of the Reference Model: As the previous stage is often fully automated and, even if not, limited by the contents of the input model data, the acquired reference model should be manually postprocessed in this stage. This may include

connecting unconnected model parts, adding, deleting or renaming nodes, or complementing the reference model by deductively developed model parts.

6. Evaluation and Enhancement of the Reference Model: A reference model will be especially useful if it is regularly evaluated and replenished as necessary. New individual models have to be included, as they arise. If the process domain changes significantly, it might make sense to develop a completely new reference model as well.

3 Identification and Analysis of Relevant Literature

3.1 Literature Review

In order to identify practical challenges to the inductive development of reference models, the first step is to identify relevant contributions to the field, i.e. literature that possibly describes such challenges. Therefore, we conduct a literature review, following the methodology as proposed by Fettke [Fe06]. After defining the objective and research question of our literature review, we conduct the literature search by entering pre-defined keywords in several research databases. The obtained results are assessed regarding their relevance to the posed research question. Thereby, we focus on contributions that concretely describe an inductive method, as opposed to methodically open ones (e.g. [Th06]). While those contributions may generally apply to the inductive strategy, they do not report on any concrete challenges, hence not providing answers relevant to our research question. In this assessment, we follow the Model of System Design [FHL10]: The respective contribution has to describe a technique, i.e. a reliable means to realize a certain goal, and explicate its foundation on empirical knowledge in order to be classified as inductive.

For conducting a literature review, a set of keywords has to be defined. Since the term "Inductive Reference Model Development" is not consistently used, several different keyword combinations assure that as much literature as possible is covered. We combine the term "reference model" with different verbs describing the process of automatic model development. Each verb can appear in either its infinitive form ("develop"), as a gerund ("developing"), or as the corresponding noun ("development"). Overall, there are eight keyword combinations, as listed in Table 1. The search is limited to title, keywords, and abstract of the articles. If those indicate a relation to reference modeling the full text is assessed. We conduct the search in five databases, namely Academic Search Complete (ASC, via EBSCO)[2], AIS Electronic Library (AISeL)[3], Business Source Premier (BSP, via EBSCO)[4], io-port [5], and Scopus[6].

After accumulating all relevant literature, we conduct a backward search, analyzing the citations of the identified articles to account for prior contributions to consider. Finally,

[2] http://search.ebscohost.com/
[3] http://aisel.aisnet.org/
[4] http://search.ebscohost.com/
[5] http://www.io-port.net
[6] http://www.scopus.com

we use *Google Scholar* to conduct a forward search, finding articles that cite the articles identified in the previous steps. As reference modeling is traditionally a topic of the German-speaking research on information systems, we also conduct an equivalent literature search using German keyword combinations. As the chosen databases mainly contain English articles, this search does not yield any significant results. However, we include German-language contributions in our forwards and backwards search to yield a better coverage of the literature.

Search Terms	ASC		AISeL		BSP		io-port		Scopus	
	All	Rel.	All	Rel.	All	Rel.	All	Rel.	All	Rel.
“reference model” AND inductive	1	0	7	2	0	0	0	0	11	2
“reference model” AND mining	10	1	8	0	8	1	19	4	163	3
“reference model” AND automatic	25	0	3	0	8	0	0	0	391	0
“reference model” AND (generation OR generate OR generating)	118	1	25	0	45	1	43	3	920	5
“reference model” AND (construction OR construct OR constructing)	72	0	52	0	30	0	45	1	692	0
“reference model” AND (development OR develop OR developing)	278	0	188	2	194	0	137	0	2859	4
“reference model” AND (derivation OR derive OR deriving)	59	0	32	1	19	0	26	1	355	0
“reference model” AND (discovery OR discover OR discovering)	38	1	2	0	20	0	18	1	142	3

Tab. 1: Quantitative Results of the Literature Review

The quantitative results of the literature search are displayed in Table 1. For each keyword combination and each database, we list the overall number of search results (in column “All”) and the number of relevant search results (in column “Rel.”). The relevance of an article is assessed manually, based on its title, keywords, and abstract. While the number of search results differs considerably, the number of relevant results is consistently very small. The large number of insignificant results is caused by the fact that, although relevant to the field of reference modeling, many of the used search terms (e.g. “construction” or “generation”) can be used in completely different contexts. In addition, many research disciplines use reference models as a scientific tool or methodology.

Accumulating all relevant articles and eliminating duplicates results in a total of ten contributions that we consider relevant for our research question. Performing a forwards and backwards search yields an additional eight articles, resulting in a total of 18 contributions, as listed in table 2. Of these 18, three [AFF11, GS14, KPR08] have the character of a practical experience report on inductive reference modeling, whereas the remaining 15 take on a rather scientific point of view.

3.2 Analysis of Contributions

In order to identify challenges to inductive reference model development, we analyze each contribution for challenges that are either explicitly mentioned as such, or implicitly addressed by according measures. For example, a paper may require previously defined activity correspondences in form of identical descriptions (implicit), while others use dedicated techniques for correspondence identification (explicit). The results of this assessment are shown in table 2. Explicit references are symbolized by a full circle (●), implicit references by a half-full circle (◐). An empty circle (○) means that the respective challenge is not mentioned in this contribution. All challenge have to be explicitly mentioned at least once. In order to provide a better overview and to point out relations, each challenge is allocated to one of the six stages of reference modeling. An in-depth description of each challenge is contained in the following section 4.

	Challenge	[AFF11]	[Ar13]	[GvJV08]	[GS14]	[KPR08]	[LRW09]	[LRW10]	[LRW11]	[MFL14]	[MFL15]	[MG13]	[PIŠK12]	[RFL13]	[RFL15]	[WFL13]	[YB11]	[Ya12]	[YWB12]
1	Indication Criteria	○	○	○	○	○	○	○	●	○	○	○	●	○	○	●	○	○	○
	Modeling Language	○	●	◐	●	●	●	●	●	●	●	◐	●	○	●	○	○	○	○
	Modeling Conventions	○	●	◐	●	●	●	●	●	○	◐	◐	○	○	●	●	●	●	●
2	Model Character	◐	○	○	◐	●	○	○	○	●	○	○	●	○	●	○	○	○	○
	Model Access	◐	○	●	◐	●	○	◐	◐	○	○	○	◐	○	●	○	○	○	○
	Collection Effort	◐	○	◐	◐	◐	○	◐	◐	○	○	○	◐	○	●	○	○	○	○
3	Modeling Language	◐	○	○	○	◐	◐	◐	◐	○	○	○	○	○	●	○	◐	◐	◐
	Modeling Conventions	◐	○	○	○	◐	◐	◐	◐	○	○	○	○	○	○	●	◐	◐	◐
	Correspondences	●	●	●	●	●	○	◐	●	●	●	○	●	○	●	◐	○	○	○
4	Abstraction	○	●	○	◐	●	○	○	○	◐	○	○	●	●	●	●	○	○	○
	Model Clustering	●	○	○	◐	○	○	○	○	○	●	○	○	○	●	●	○	○	○
	Decomposition	○	○	○	●	●	○	○	○	○	○	○	◐	○	○	○	○	○	○
	Model Variants	◐	○	●	○	●	◐	●	●	○	○	○	◐	○	●	○	○	○	○
	Contortion	○	○	○	○	○	○	○	○	●	○	○	○	○	●	○	○	○	○
	Inductive Fallacy	●	○	○	○	○	○	○	○	○	○	○	○	●	○	○	○	○	○
	Complexity	○	●	◐	○	○	●	●	●	●	●	○	○	○	●	○	●	●	●
5	Additions	○	○	○	◐	●	○	○	○	○	○	○	○	○	●	●	○	○	○
	Model Connection	○	○	○	◐	●	○	○	○	○	◐	○	○	○	●	●	○	○	○
6	Evaluation	●	○	○	◐	●	○	○	○	○	○	○	●	○	○	○	○	○	○
	New Models	○	○	○	◐	○	○	○	●	○	○	○	○	○	○	○	○	○	○

Tab. 2: Foundation of Identified Challenges in Literature

The results of our assessment allow for several interesting insights. First of all, none of the challenges is particular to one specific contribution. Each of them is mentioned at least twice, although sometimes implicitly, meaning that the respective contribution describes measures to address this challenge, without explicating why or how it influences the reference model development. Most of these implicit references appear in case studies or empirical evaluations, such as [AFF11] or [GS14]. Rather epistemological challenges, such as the problem of input data contortion or inductive fallacies are rarely mentioned. Instead, most contributions focus on practical or technical challenges, such as reaching a feasible algorithmic complexity or ways to establish correspondences between individual models in terms of a matching. While these challenges are unquestionably important for the practical application of the inductive strategy, epistemological aspects should also be considered, as they play an important role for evaluating the resulting reference model.

Regarding the six stages of reference modeling, it is apparent that stage four, concerned with the derivation of the reference model, contains the most challenges and is addressed by every single contribution. As this stage contains the inductive reference model development in a narrower sense, it is the core of most inductive development approaches and hence poses more challenges that have to be addressed. The preparatory first stage is also well-covered, although this is mostly caused by assumptions and restrictions to the input data. Nevertheless, this critical assessment is important for developing new approaches to reference modeling, which can potentially overcome these restrictions. On the contrary, both the collection stage (two) and the postprocessing and evaluation stages (five and six) are only sparsely covered by the contributions. While it is possible to develop a reference model without postprocessing and evaluating it, the collection of the input models is central to the success of the model development itself, as the contents of the reference model is directly influenced by the contents and character of the input models. It is remarkable that the challenges of this stage are only addressed by about half the contributions and mostly in an implicit way. Many contributions assume to have access to a sufficiently large and representative collection of digitally represented individual models, which simplifies most research activities, but is a non-trivial problem in practical applications.

Regarding the contributions, it is apparent that some take on a much broader view on inductive reference model development than others. Especially the practical reports on inductive reference model development [AFF11, GS14, KPR08] cover many challenges that are not considered by other, rather methodical contributions. Among these, the degree of examining potential challenges differs considerably. While some contributions try to take a rather hollstic view on reference model development [RFL15, LRW11], others focus on one particular challenge, which is analyzed in depth (for example input model abstraction in [RFL13]). Other contributions have a rather technical view on reference model development, focusing on algorithmic aspects and technical details (e.g. [MG13, YB11]).

4 Current Challenges of Inductive Reference Model Development

4.1 Preparation of Reference Model Development

Before beginning to collect data, several preliminary questions have to be answered. Essential challenges are:

- Indication Criteria: It is indisputable that according individual models have to be available to enable inductive reference model development. If, during the preparatory stage, it is unclear whether an adequate amount of individual models exist or can be easily acquired, the entire project is at risk. Hence, we require criteria determining when to resort to an inductive development and when to avoid it. A well-founded empirical knowledge base is a crucial success factor for the inductive strategy [WFL13]. Moreover, for organizational reasons, the development of a new reference model might not always be supported [LRW11].

- Choice of Modeling Language: The choice of the reference modeling language in the context of the inductive strategy is important for two reasons. First of all, it makes sense to follow the individual models in order to avoid unnecessary conversions and transformations. Therefore, algorithmic approaches as described in Martens et al. [MFL14] or Yahya et al. [YWB12] require a formal mostly graph-based representation of the individual models and the resulting reference model. Second, due to the number and variety of the individual models to process, it is often necessary to represent the model variants in an appropriate way. Hence, the target language should provide concepts for variant management as described in [GvJV08].

- Modeling Conventions: In general, modeling conventions are regarded to be important. However, due to the plurality of modeling contexts, it cannot be assumed that modeling conventions are always factually enforced when creating the input models. Hence, according compromises have to be found. Especially the algorithmic solutions we identified address this challenge, since they rely on characteristics such as unique labels [LRW09] or block-structuredness of the input process models [Ar13].

4.2 Collection of Individual Models

Individual enterprise models are an obligatory input and thus a necessary requirement for applying the inductive development strategy. They can either be collected originary (primary collection)[PIŠK12] or existing information can be reused (secondary collection) [GvJV08, AFF11]. Essential challenges are:

- Representative Character of the Individual Models: If the representative character of the individual models is questionable, there is a risk of contortion. For instance, particular model variants may be overrated in the derivation of the reference model. Meanwhile, recognizing such contortions requires to know the entire model domain. In some cases, however, this domain can only reasonably be assessed after the reference model has been described. In other applications, it may make sense to purposefully avoid a representative character, for instance to examine contrasting individual models. In this case, the question arises how many contrasting cases are required. Hence, Karow et al. [KPR08] require a rich empirical database and an inductive validation of the reference model. Pajk et al. [PIŠK12] address this challenge by first developing the model on the empirical data available and then assessing it according to predefined requirements.

- Access to Individual Models: According to Karow et al. [KPR08], accessing enterprise models can be difficult to impossible. Individual models can be subject to confidentiality or organizations see the risk to loose competitive advantages by allowing them to circulate. Hence, cooperation cannot be expected when applying the inductive development strategy. Therefore, the proof of concept for the method proposed in [KPR08] is settled in the field of public administration and it is based on public domain models.
- Collection Effort: The effort necessary for collecting the individual models can overcompensate the benefits of reference modeling. This risk is especially apparent, if cost and benefits are very differently distributed among the stakeholders [RFL15].

4.3 Preprocessing of Individual Models

Before concrete methods for reference model acquisition can be applied, a general preprocessing of the individual models appears to be useful. Essential challenges are:

- Transformation of the Modeling Language: If the individual models are not represented in a uniform modeling language, it is suitable to transform and convert them into a uniform target language. This is crucial for the algorithmic approaches we obtained, since the processing relies on a uniform representation, for example as an Event-Driven Process Chain [RFL15]
- Examination and Establishment of Modeling Conventions: If certain conventions have been determined, it makes sense to examine the models for meeting them. If deviations exist, according corrections of the individual models are required. It may also make sense to reconstruct the factually used conventions [WFL13].
- Identification of Correspondences: Single constructs in different individual models can be equivalent, similar, or different. It is essential to know about such correspondences. The need for and the challenges of identifying these correspondences is addressed in most of the approaches that aim at an algorithmic solution [MFL15, Ar13, RFL15]. The correspondences can be derived from solely structural characteristics of the input models (for example similar graph-theoretic model features [RFL15]) or be related to semantic aspects (for example similar process traces or words in node labels [Ar13, MFL15, MFL14]). The fact that structural similarities are neither necessary nor sufficient for semantic similarities, is an additional challenge.

4.4 Acquisition of the Reference Model

The aim of the reference model acquisition stage is to derive a reference model from a set of individual models. Essential challenges are:

- Abstraction and Generalization: Methods for abstracting specific model features are required to generalize the contents of the individual models. If the individual models contain commonalities, they are to be represented in the reference model. This way, the reference model contains the typical structures of the application domain. Hence, the acquired reference model should abstract from specific features of individual models. In analogy to the identification of correspondences described above, not only the model structure, but also the model semantics should be considered. In [RFL13], a variety of abstraction approaches are evaluated in terms of applying them in the context of inductive reference model development.
- Clustering of Individual Models: Developing model clusters is important for two reasons. First, individual models can stem from completely different contexts, so a cluster analysis can be used to identify possible domains [WFL13](for instance for cross-sectional tasks like accounting or logistics). Second, Aier et al. [AFF11] postulate a cluster analysis in order to group a variety of model variants within the same domain.
- Decomposition of Individual Models: For the decomposition, two aspects should be considered. First, individual models often consist of numerous partial models, which should be analyzed separately [GS14]). Second, even a single partial model may be decomposed into fragments. This makes sense, because two models can be different, while individual model fragments may be similar or even identical.
- Plurality of Model Variants: When acquiring a reference model, two opposing design objectives have to be balanced. One extreme case includes all individual models as model variants in the input model, the other extreme case only captures the commonalities of the individual models in the reference model and abstracts from individual model features. Aier et al. [AFF11] describe this challenge as a fundamental conflict between standardization and individualization of the reference model.
- Handling Contortions: If the collected individual models do not constitute a representative sample of the domain, it is to be checked whether possible contortions exist and can be corrected by taking according measures. Contortions can be prevented by selecting a representative or sufficiently large set of organizations [RFL15].
- Risk of Inductive Fallacy: Philosophy in general and (inductive) logic in particular often refer to the weaknesses of inductive conclusions [AFF11]. This aspect should also be considered in the inductive development strategy. The fact that models may contain descriptive as well as prescriptive content also has to be taken into account [RFL13].
- Algorithmic Complexity: There exists a variety of algorithms for model analysis that include techniques for comparing process models. These techniques include graph-based structural comparisons [YB11, Ya12, YWB12] as well as comparisons that deal with the processessing and comparison of natural language [Ar13, MFL15]. Due to their high runtime and memory consumption, such algorithms are unable to cope with larger model sets as they typically appear within the inductive development strategy.

4.5 Postprocessing of the Reference Model

Essential challenges when postprocessing a reference model are:

- Additions: The inductive approach of Grger and Schumann [GS14] includes an enhancement of the inductively derived reference model. Reference models can be expected to be complemented by additional contents. The question arises, how to combine inductive and deductive strategies into a hybrid development strategy.
- Connection of Partial Models: It should be considered to group separate reference models. Grouping makes sense, if the individual models were accordingly decomposed in a previous step [GS14, KPR08]. Adding such connections is also useful to point out similarities between reference models of different domains (for example warehousing processes in retail and industry).

4.6 Enhancement of the Reference Model

Reference modeling is especially useful if the reference model is constantly evaluated and enhanced. Otherwise, there is a risk of obsolescence. Essential challenges are:

- Evaluation: In this context, it is important to assess the generated reference model with regard to the predefined goals. Dependening on the specific stakeholder or perspective, the verifciation and validation varies. While Aier et al. [AFF11] conduct interviews with process users, Pajk et al. [PIŠK12] propose a simulation or even an empirical evaluation by applying the reference model.
- New Individual Models: If new variants are derived from a reference model, they should be considered to be included in the reference model. Reference model stability should be aimed for, i.e. not every variant should be included. However, it is obviously sensible to include standard adaptations, which are regularly executed during an application, into the reference model. It might be desirable to constantly enhance the reference model to minimize its distance to all existing variants [LRW11].

5 Discussion and Conclusion

The objective of the article at hand was to identify a catalog of potential challenges to the practical application of the inductive development strategy for reference modeling. This catalog is meant to stimulate further research by providing a list of empirically grounded challenges, which can be addressed in a structured way to advance the inductive strategy both in research and practice. However, we do not claim our list of challenges to be complete or exhaustive. For example, none of the contributions discusses the possibility for a regular re-design of the reference model, which could become necessary, if the input models change considerably. In future research, we intend to validate and extend the identified challenges by conducting expert interviews before dealing with each challenge individually. Interviewing reference modeling experts can offer a complementary perspective and lead to additional insights from a practitioner's point of view.

Our contribution in this article can only be a starting point for the in-depth analysis of the challenges to inductive reference modeling. The term "reference model" is mainly popular in German-speaking countries and Australia, while other communities prefer terms such as "reusable model". Extending our literature review by those terms can further advance our findings. Meanwhile, we cannot assure that additional challenges are mentioned in previous works on reference modeling, not focusing on the inductive strategy. However, it is almost impossible to assess their relevance without taking explicit features of the inductive strategy, such as the foundation on empirical knowledge, into account.

It should also be noted that none of the identified contributions contains a direct reference towards Business Process Management in particular, although it appears self-evident that inductive reference modeling offers interesting potentials for classical BPM topics, such as continuous process improvement, customization and implementation of standard software, or process documentation. Traditional applications of reference modeling, such as design, optimization, simulation, or certification of business processes may also be addressed by inductive methods. Simultaneously, many identified challenges are related to topics that are already addressed by BPM researchers. For example, advancements on naming conventions, model enhancement, or variant management could be built upon to deal with the according challenges.

Despite the numerous challenges presented in this article, the inductive development strategy has considerable advantages when compared to the deductive strategy. Analyzing individual models and their potentials for generalization is not only intellectually interesting, but appears almost obligatory if the models are easily available. Second, a direct comparison does not require general principles and theories to derive a reference model. Moreover, an inductively developed reference model can be expected to have a higher degree of detailing, maturity, and acceptance. It thus appears to be especially fruitful to advance the integration of inductive and deductive development into a hybrid strategy. In general, inductive reference modeling is still at a rather early stage of development and there is a lot of potential for further research.

In addition, methods of inductive reference modeling offer interesting potentials for related applications. For example, model comparison, model integration, or analysis of internal model variants come to mind. Besides these practical applications, there are also interesting potentials for an application in research. Inductive methods can for example be used for a scientifically well-grounded comparison of models. They are hence an important foundation for one of the Grand Challenges identified by Mertens and Barbian [MB15]: Following the Human Genome Project, the "Wirtschaftsinformatik" should chart existing IT systems in order to draw conclusions for the future employment of IT and the accuracy of its systems. Such activities may significantly benefit from inductive methods and thus discover new potentials for known works on organizational typology and the model of shell and nucleus.

References

[AFF11] Aier, Stephan; Fichter, Michael; Fischer, Christian: Referenzprozesse empirisch bestimmen : von Vielfalt zu Standards. Wirtschaftsinformatik & Management : WuM, 3(3), 2011.

[Ar13] Ardalani, Peyman; Houy, Constantin; Fettke, Peter; Loos, Peter: Towards a Minimal Cost of Change Approach for Inductive Reference Model Development. In: Proceedings of the 21st European Conference on Information Systems. European Conference on Information Systems (ECIS-13). AIS, 2013.

[Be02] Becker, Jörg; Delfmann, Patrick; Knackstedt, Ralf; Kuropka, Dominik: Konfigurative Referenzmodellierung. In (Becker, J.; Knackstedt, R., eds): Wissensmanagement mit Referenzmodellen: Konzepte für die Anwendungssystem- und Organisationsgestaltung, pp. 25–144. Springer, 2002.

[BS97] Becker, Jörg; Schütte, Reinhard: Reference Information Systems for Retail: Definition, Use and Recommendations for Design and company-specific Adaption of Reference Models. In: Wirtschaftsinformatik, pp. 427–448. Springer, 1997. In German.

[DM05] Daun, Christine; Matheis, Thomas: Constructing a reference process model for E-government. In (Mosca, R., ed.): Proceedings of the 7th International Conference on The Modern Information Technology in the Innovation Processes of the Industrial Enterprises (MITIP), Genua. pp. 10–14, 2005.

[Fe06] Fettke, Peter: State-of-the-Art des State-of-the-Art. Wirtschaftsinformatik, 48(4):257–266, 2006.

[Fe14] Fettke, Peter: Eine Methode zur induktiven Entwicklung von Referenzmodellen. In (Kundisch, Dennis; Suhl, Leena, eds): Tagungsband Multikonferenz Wirtschaftsinformatik 2014. Universität Paderborn, 2014.

[FHL10] Fettke, Peter; Houy, Constantin; Loos, Peter: On the relevance of design knowledge for design-oriented business and information systems engineering. Business & Information Systems Engineering, 2(6):347–358, 2010.

[FL07] Fettke, Peter; Loos, Peter: Perspectives on Reference Modeling. In (Fettke, Peter; Loos, Peter, eds): Reference Modeling for Business Systems Analysis, pp. 1–20. Idea Group Publishing, 2007.

[GS14] Gröger, Stefan; Schumann, Matthias: Entwicklung eines Referenzmodells für die Gestaltung des Drittmittel-Prozesses einer Hochschule und Ableitung von Einsatzgebieten für Dokumenten-und Workflow-Management-Systeme. Arbeitsbericht der Universität Göttingen, 1, 2014.

[GvJV08] Gottschalk, Florian; van der Aalst, Wil; Jansen-Vullers, Monique: Mining Reference Process Models and Their Configurations. In: On the Move to Meaningful Internet Systems: OTM 2008 Workshops. Springer, pp. 263–272, 2008.

[KPR08] Karow, Milan; Pfeiffer, Daniel; Räckers, Michael: Empirical-Based Construction of Reference Models in Public Administrations. In: Multikonferenz Wirtschaftsinformatik. pp. 1613–1624, 2008.

[LRW09] Li, Chen; Reichert, Manfred; Wombacher, Andreas: Discovering Reference Models by Mining Process Variants Using a Heuristic Approach. In (Dayal, Umeshwar; Eder, Johann; Koehler, Jana; Reijers, Hajo, eds): Business Process Management, volume 5701 of Lecture Notes in Computer Science, pp. 344–362. Springer Berlin Heidelberg, 2009.

[LRW10] Li, Chen; Reichert, Manfred; Wombacher, Andreas: The MinAdept clustering approach for discovering reference process models out of process variants. International Journal of Cooperative Information Systems, 19(03n04):159–203, 2010.

[LRW11] Li, Chen; Reichert, Manfred; Wombacher, Andreas: Mining business process variants: Challenges, scenarios, algorithms. Data & Knowledge Engineering, 70(5):409–434, 2011.

[MB15] Mertens, Peter; Barbian, Dina: Researching "Grand Challenges". Business & Information Systems Engineering, 57(6):391–403, 2015.

[MFL14] Martens, Alexander; Fettke, Peter; Loos, Peter: A Genetic Algorithm for the Inductive Derivation of Reference Models Using Minimal Graph-Edit Distance Applied to Real-World Business Process Data. In (Kundisch, Dennis; Suhl, Leena; Beckmann, Lars, eds): Tagungsband Multikonferenz Wirtschaftsinformatik 2014. Universität Paderborn, 2014.

[MFL15] Martens, Alexander; Fettke, Peter; Loos, Peter: Inductive Development of Reference Models Based on Factor Analysis. In (Thomas, Oliver; Teuteberg, Frank, eds): Proceedings der 12. Internationalen Tagung Wirtschaftsinformatik (WI 2015). volume 12, Universität Osnabrück, Osnabrck, Germany, pp. 438 – 452, 2015.

[MG13] Mejri, Asma; Ghannouchi, Sonia Ayachi: Discovering Reference Process Models in the Context of BPM Projects. Procedia Technology, 9:489–497, 2013.

[PIŠK12] Pajk, Dejan; Indihar-Štemberger, Mojca; Kovačič, Andrej: Reference model design: An approach and its application. In: Information Technology Interfaces (ITI), Proceedings of the ITI 2012 34th International Conference on. IEEE, pp. 455–460, 2012.

[RFL13] Rehse, Jana-Rebecca; Fettke, Peter; Loos, Peter: Eine Untersuchung der Potentiale automatisierter Abstraktionsansätze für Geschäftsprozessmodelle im Hinblick auf die induktive Entwicklung von Referenzprozessmodellen. In (Alt, Rainer; Franczyk, Bogdan, eds): Proceedings of the 11th International Conference on Wirtschaftsinformatik, Leipzig, Germany. 2013. In German.

[RFL15] Rehse, Jana-Rebecca; Fettke, Peter; Loos, Peter: A graph-theoretic method for the inductive development of reference process models. Software & Systems Modeling, tba:1–41, 2015.

[Th06] Thomas, Oliver: Management von Referenzmodellen: Entwurf und Realisierung eines Informationssystems zur Entwicklung und Anwendung von Referenzmodellen. Logos-Verlag, 2006.

[WFL13] Walter, Jürgen; Fettke, Peter; Loos, Peter: How to Identify and Design Successful Business Process Models: An Inductive Method. In (Becker, Jörg; Matzner, Martin, eds): Promoting Business Process Management Excellence in Russia - Proceedings and Report of the PropelleR 2012 Workshop. volume 15. European Research Center for Information Systems, Münster, pp. 89–96, 2013.

[Ya12] Yahya, Bernardo; Bae, Hyerim; Bae, Joonsoo; Kim, Dongsoo: Generating valid reference business process model using genetic algorithm. International Journal of Innovative Computing, Information and Control, 8(2):1463–1477, 2012.

[YB11] Yahya, Bernardo; Bae, Hyerim: Generating Reference Business Process Model Using Heuristic Approach Based on Activity Proximity. In: Intelligent Decision Technologies, pp. 469–478. Springer, 2011.

[YWB12] Yahya, Bernardo; Wu, Jei-Zheng; Bae, Hyerim: Generation of Business Process Reference Model Considering Multiple Objectives. Industrial Engineeering & Management Systems, 11(3):233–240, 2012.