

Quo Vadis Modeling?

Findings of a Community Survey, an Ad-hoc Bibliometric Analysis, and Expert Interviews on Data, Process, and Software Modeling

Judith Michael · Dominik Bork · Manuel Wimmer · Heinrich C. Mayr

Received: date / Accepted: date

Abstract Models are the key tools humans use to manage complexity in description, development, and analysis. This applies to all scientific and engineering disciplines, and in particular to the development of software and data-intensive systems. However, different methods and terminologies have become established in the individual disciplines, even in the sub-fields of Informatics, which raises the need for a comprehensive and cross-sectional analysis of the past, present, and future of modeling research. This paper aims to shed some light on how different modeling disciplines emerged and what characterizes them with a discussion of the potential toward a common modeling future. It focuses on the areas of software, data, and process modeling and reports on an analysis of the research approaches, goals, and visions pursued in each, as well as the methods used. This analysis is based on the results of a survey conducted in the communities concerned, on a bibliometric study, and on interviews with a prominent representative of each of these communities. The paper discusses the different viewpoints of the communities, their commonalities and differences, and identifies possible starting points for further collaboration. It further discusses current challenges for the communities in general and modeling as a research topic in particular and highlights visions for the future.

Keywords Research Communities · Software Engineering · Software Modeling · Data Modeling · Process Modeling · Information Systems

1 Introduction

Motivation. In November 2018, 36 researchers from 15 countries met for a "Dagstuhl Seminar" at the Leibniz

J. Michael
RWTH Aachen
E-mail: michael@se-rwth.de

D. Bork
Business Informatics Group, TU Wien, Austria

M. Wimmer
CDL-MINT, Johannes Kepler University Linz, Austria

H.C. Mayr, Alpen-Adria Universität Klagenfurt, Austria

Center for Informatics (Wadern, Germany)¹. The title of the seminar was "Next Generation Domain-Specific Conceptual Modeling: Principles and Methods"². The organizers had intentionally placed the focus on domain-specific methods, as these are assumed to be more tailored to the needs of the particular user group or community and are less burdened with foundational and cross-community issues.

However, already on the first day, it became clear that even within this narrow scope, discussions about terminology, methods, and fundamentals had to be held. The participants came from different sub-disciplines of informatics in which modeling has an important position, e.g., software and systems engineering, database engineering, and business informatics. We were surprised by the wide variety of different views on, e.g., the term "conceptual modeling" and the notion of a "model" (cf. [71]).

Due to the importance of conceptual modeling, there have also been recent efforts to establish robust theoretical foundations for this field [27, 55, 62, 70]. As these efforts are not yet finished and we are more interested in methods than in terminology, we have decided to take an observer role and simply speak of "modeling" in this paper. This is also aligned with the notion of a recent paper that raises the demand that "Modeling should be an independent scientific discipline" [17]. The authors argue, that the rich modeling expertise in the field of software engineering should be transferred to other scientific fields and that, conversely, the software community could benefit from such cooperation with other communities—a viewpoint with which we fully agree.

¹ www.dagstuhl.de

² www.dagstuhl.de/de/programm/kalender/semhp/?semnr=18471

Contribution. Our goal is to get insights into the various modeling communities, their topics and visions, and the foundations, methods, and terminologies they use. We have limited ourselves to looking at only three communities in more detail, namely *Software Modeling*, *Data Modeling*, and *Process Modeling*. A first look may convey the impression that there seems to be little exchange between these communities, although they overlap slightly. For instance, they publish in different outlets (i.e., conferences and journals) and attend at different conferences, however, some topics are reoccurring in these communities. This makes it difficult, for example, for researchers, and especially, PhD students, to spread their work as widely as possible or to switch between communities during their careers. Motivated by these first insights and the discussion at the Dagstuhl seminar, we aim to contribute to the improvement of this situation by conducting a systematic analysis of the state of the modeling research across and within the three modeling communities of data, process, and software modeling. With the study reported here, we aim to provide answers to the following questions:

1 Research Topics

- Q1.1 Which were, are, or will be the main research topics and application areas?
- Q1.2 Which are the main foundations?
- Q1.3 Which are the main methodologies?
- Q1.4 Are there differences between research and practice with respect to research topics and application areas?
- Q1.5 What does modeling have to achieve to increase its importance in 10 years' time?

2 Exchange across communities and between research and industry

- Q2.1 How often do researchers publish in different communities?
- Q2.2 What are the community-specific and the community-spanning research topics?
- Q2.3 How much cooperation do researchers want across modeling communities?
- Q2.4 What is the state of cooperation between practice and research and what do they expect from each other?

In formulating these questions, we have taken into account aspects of the *past*, the *present*, and the *future*. We used a mixed method approach that combined qualitative methods, e.g., interviews, and quantitative methods, e.g., data analysis, to respond to these questions (see Section 3).

Outline. The rest of this paper is structured as follows. In Section 2 we present a conceptualization of the notion of a research community on which we can build

further considerations. Section 3 describes our framework for analyzing the communities. Section 4 discusses our main results focusing on the past and the present modeling research while Section 5 moves the focus toward the future visions for modeling. Section 6 contains summaries of the transcribed interviews we conducted with Jean Bézivin, Peter P.S. Chen, and Wil van der Aalst. Section 7 discusses the results before we summarize and conclude our paper in Section 8.

2 A Notion of “Research Community”

Models are working instruments for nearly all scientific, engineering, and application-oriented domains, *e.g.*, in medicine to understand the human body, in architecture to design or redesign buildings and objects, or in sports science to analyze and improve athletic performance [72]. Modeling thus comes along with great diversity which is manifested in, *e.g.*, (1) the various disciplines and application areas, (2) the objectives for which modeling is used therein, and (3) the methods employed in each case.

This also applies to the field of Informatics, where models are used for a wide variety of purposes, *e.g.*, for database design, analyzing, simulating, documenting, refactoring, rapid prototyping, testing, or (iterative) code generation [43]. It is therefore not surprising that different communities have emerged over time, each publishing their results in different publication outlets and attending different conferences to exchange ideas, without, of course, being completely disjoint. Thus, if we are to examine the research topics, foundations, and methods of such communities and their “exchange” in more detail, we must first sharpen the notion of community itself, limiting ourselves here to the conceptualization of the term “*research community*”.

First, we found a number of literature sources in which (research) communities are treated. For example, initiatives aimed at establishing a Body of Knowledge for a specific scientific discipline (or community), [11, 14, 44, 76] should be mentioned here. Also, work can be found on analyzing ‘research communities’ of a specific discipline [16, 47] or within geographic areas [24, 37]. Previous research further focused on introducing metrics for community assessment [21] or techniques and tools [83] for the analysis of, *e.g.*, the topics a community is interested in [54], its contributors [53, 84], the research methods used [35], or a combination of them within a scientific community [59]. Most of these approaches are based on bibliometric data or paper full texts on which algorithms are applied to compute community metrics like the most active and influential authors, co-authorship networks, the closeness of a

community, and many more. For deriving these metrics most approaches rely on (a combination of) techniques spanning from conventional analytics (e.g., algorithms applied to CSV or Bibtext files), over network analysis and graph analysis techniques up to the application of Neural Networks for the prediction of research trends [52].

What is missing, however, is an attempt to define a notion of “research community” and a more in-depth analysis of the data, process, and software modeling communities we are focusing on. Therefore, in the following, we introduce our attempt to derive a first coarse-grained notion of a research community by means of a conceptual model shown in Figure 1. We omit attributes of the individual entity and relation sets since we are only interested in the essential concepts of a research community and its relationships.

According to this model, a **Research Community** has *Researchers*, who deliver *Contributions* (*Papers* or *Services*) to *Research Areas* that are *addressed* by the Research Community. These contributions are made to *Platforms* (*Conferences* or *Journals*) *used* or operated by the research community. Note that *Conference* is understood very broadly here and includes events such as symposia, workshops, etc. *Paper* stands for articles, essays, posters (recorded) lectures etc., and typical *Services* include involvement in editorial boards, program committees, steering committees, etc.

The only additional (and reflexive) relationship we have included is the citation relationship between papers since we believe that clustering can also be derived from this in practice.

We deliberately do not impose strong multiplicity restrictions (such as the limitation to one research area) in order to allow for an easy adaptation of the model to particular cases of investigation. Consequently, it is possible, among other things, that:

- a Researcher belongs to several Research Communities,
- a Communication Platform is used by several Research Communities, or
- a particular Paper is cited by Contributions to different Research Areas.

When instantiating this model in the scope of this study, we, first of all, choose *Modeling* as the *Research Community*. The further subdivision within this community is then done via the *Communication Platforms*, i.e., *Conferences* and *Journals*. It was obvious for us to consider the following conferences as the premier venues for the formation of the respective communities:

- Data modeling: *International Conference on Conceptual Modeling (ER)*; held annually since 1979.

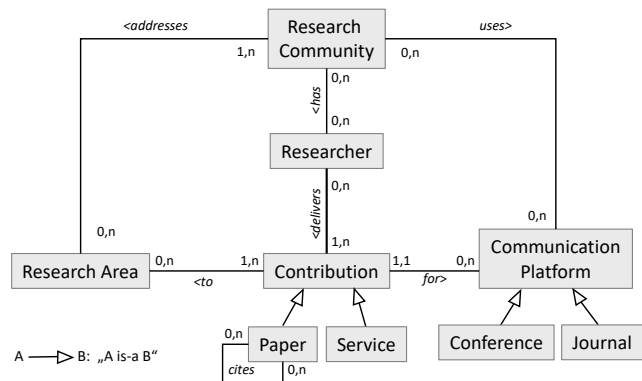


Fig. 1 Conceptual model of “research community” (multiplicities in ER reading style)

- Software modeling: *International Conference on Model-Driven Engineering Languages and Systems (MODELS)*; held annually under this name since 2005, was founded in 1998 as the International Conference on The Unified Modeling Language (UML).
- Process modeling: *International Conference on Business Process Management (BPM)*; held annually since 2003.

We admit, that many other conferences with an account for data, software, and process modeling exist, but we decided to focus on the one, premier outlet to best represent the community. As these are also amongst the most selective outlets with respect to the acceptance rate and the most prestigious to get accepted, we believe they best represent the state of affairs of the respective modeling community. The assignment of journals which are specific to only one modeling community is further complicated, even though significant overlaps in the editorial boards of candidate journals like *Data and Knowledge Engineering (DKE)* and *Software and Systems Modeling (SoSyM)* with the steering committees of ER and MODELS conferences respectively exist. A closer look, however, shows that journals naturally are broader in scope, do not only deal with pure modeling subdivision topics, and have much broader cross-community coverage compared to the conferences.

Therefore, we decided to use the above-mentioned central conference as the first discriminator in the way that we associate an author with the respective community if the author mainly publishes at the related conference. In the further course of this paper, the acronyms of these conferences (i.e., BPM, ER, and MODELS) are used as abbreviations for the communities, which facilitates the readability of the tables in particular.

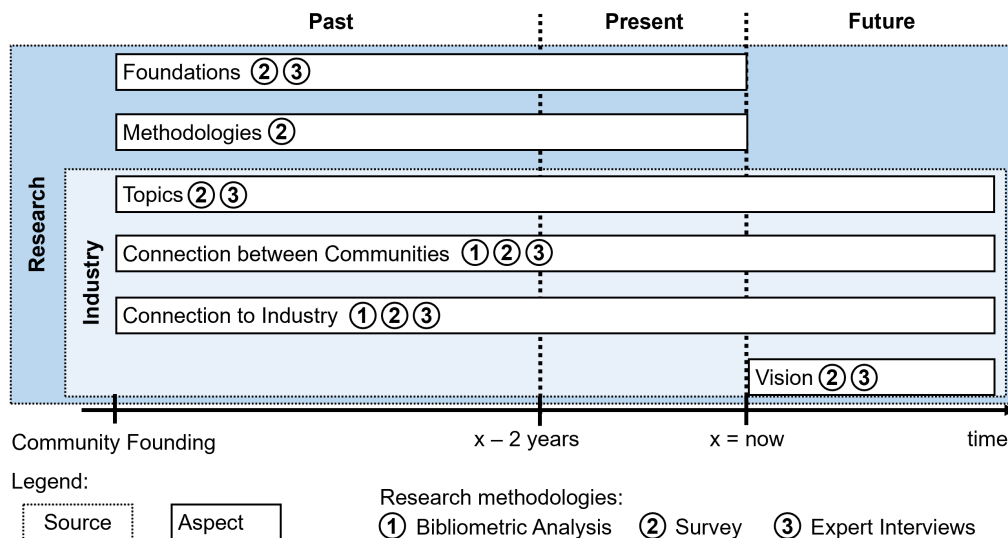


Fig. 2 Analysing Communities: The Investigated Field and Used Methods

3 A Framework for the Analysis of Research Communities

To answer the research questions afore presented, we have to define the criteria according to which we analyze the mentioned research communities in more detail. The existing literature on this is sparse, but one can find approaches that consider joint publications and cross-relationships through citations as the criteria that shape the community. For example, Sciabolazza et al. [63] apply methods from Social Network Analysis and Network Science [61, 64] to study community formation and interdisciplinary collaborations. They identify a research community as a cluster of scientists who have shared research interests, methods, and scientific approaches to problems over time and have worked together in collaborative networks. This is done on the basis of peer-reviewed publications and approved project grants. Our approach used in this paper is similar, but refined to the extent that we analyze our communities from three perspectives:

- **Aspect:** we examine various specific characteristics of a research community, namely its (a) *foundations* and (b) *methodologies*, (c) which *topics* are of interest and (d) what *visions* its members have for the future, and how closely it is (e) linked to *other research communities* and (f) *to industry*. If only one community is to be analyzed, the question of linkages between different communities would need to be omitted or specified to sub-communities within a research community. The latter approach is followed in this paper by dividing the modeling research community into three sub-communities of data, process, and software modeling.

- **Origin:** This concerns the origin (source) of the relevant information. While the aspects (a) foundations and (b) methodologies are of high relevance mainly to researchers, all other previously mentioned aspects are of interest to researchers and industry members.
- **Time:** We distinguish three periods in this context, namely (i) *present*, the time span of the past two years seen from the time of writing, (ii) *past*, the period from founding the community³ until the present, and (iii) *future*, the time from the present onwards.

Figure 2 illustrates these perspectives. In addition, it shows the three research methodologies (1) *bibliometric analysis*, (2) *survey*, and (3) *expert interviews* that we used to analyze the aspects in the ways described below. As can be seen in this figure, the bibliometric analysis helps to identify past and present aspects, and links to the industry while visions for the future can only be gathered through the survey and the expert interviews.

Bibliometric Analysis. In the *bibliometric analysis* (for a comprehensive definition see [28]), we essentially looked at the publications which are part of the premier modeling conferences aiming to analyze the collaboration between communities and the connections to the industry. We analyzed who published with whom, when, and where. The available data for

³ The real date of a community foundation is very individual depending on the community. This can range from discussions at already established conferences to workshops or seminars. In order to be able to analyze the mentioned aspects, the first year of the main conference could be used for pragmatic reasons.

this is good due to freely accessible databases. The situation with other kinds of information, such as community services, is more difficult, as there are no comprehensive databases on these yet.

We used all three premier conference proceedings of the BPM, ER, and MODELS conferences as the representative for the likewise named modeling subdivisions; each from the first year of publication until 2022. We extracted the research papers⁴ published in these three conferences from DBLP⁵ and collected measures how often a given author has published in these conferences. In total, we were able to identify 5,079 different authors, 1,364 of whom have more than one publication. Concerning the number of publications, we extracted information about 3,035 selected papers for all three conferences (1,084 for MODELS, 527 for BPM, and 1,424 for ER).

Survey. With the help of a *survey*, we aimed to obtain information from professional colleagues that is difficult or impossible to collect via bibliometric analysis. In particular, we used it to capture the assessments of industry representatives, which hardly can be derived from publications. As shown in Table 5 in the appendix, our survey was composed of five questions about the participant and 18 questions about specific assessments concerning the past, present, and future of collaborations across modeling communities and industry-research collaborations. Some questions are mandatory, others optional. Depending on the type of question, different forms of response were allowed: Selection from a Pre-set List, Free Text Fields, Likert Scales, and Preference Lists. In addition, we tailored some questions to more research- or industry-oriented participants and presented them exclusively to these “target groups” (see Table 5).

The implementation took the form of a SoSci online questionnaire⁶. To begin, a small group of people was asked to answer the first version of this questionnaire and to give us feedback (e.g., regarding the scope, type, and comprehensibility of the questions and the respective answer options). The results of this “test run” were incorporated before the questionnaire was sent to a wider audience.

The survey was sent to the participants of the Dagstuhl seminar that initiated the research of this paper, to specific mailing lists such as IS-WORLD and DB-World, and we referred to the survey at various relevant conferences. A total of 153 persons

Table 1 Descriptive statistics of the demographic data (rounded)

Variable	Category	Frequency
Gender (in percentage)	female	23,5
	male	73,2
	diverse	3,3
Age range (in percentage)	below 20	2,6
	between 20 and 29	13,7
	between 30 and 39	29,4
	between 40 and 49	20,3
	between 50 and 59	20,3
	between 60 and 69	11,8
	above 70	2,0
Job profiles (in percentage)	professors	45,1
	post docs	5,2
	scientists or lecturers	16,3
	PhD candidates	11,1
	industry positions	9,8
	retired	4,6
	teachers	0,7
	students	1,3
	other	5,9
Communities research (in percentage)	ER	42,5
	BPM	29,9
	MODELS	27,6
Communities industry (in percentage)	data	25,0
	process	33,3
	software	41,7

participated in the survey. 128 of them see themselves in research, 25 work in industrial practice.

Table 1 shows the demographic data queried about the participants. Around three-quarters of them are male, around one-quarter are female, and more than two-thirds are in the age range 30 to 59. The participants from the scientific community count themselves as 42.5% of the ER community, 29.9% of the BPM community, and 27.6% of the MODELS community. Among the industry representatives, the ratio is reversed: the largest share associates itself as belonging to the software modeling community. Only very few of them have mentioned in the survey that they have attended one of the above-mentioned conferences.

The 153 participants come from 32 countries (scientists from 31 and industry representatives from 13 countries). Figure 3 shows the countries the participants originate from. The largest number of participants were from Germany (29%) followed by Spain (12%) and Italy (9%). We categorized the single participants from Argentina, Czech Republic, Estonia, Greece, Guam, India, Iraq, Kenya, Lebanon, Poland, Romania, Ukraine, and Uruguay in the category *Other*.

Expert Interviews. Our research methodologies are concluded by expert interviews with three well-known and

⁴ We filtered keynote, journal first, tutorial, panel, and workshop summary papers as well as invited contributions.

⁵ <https://dblp.uni-trier.de>

⁶ <https://www.sosciurvey.de/en/index>

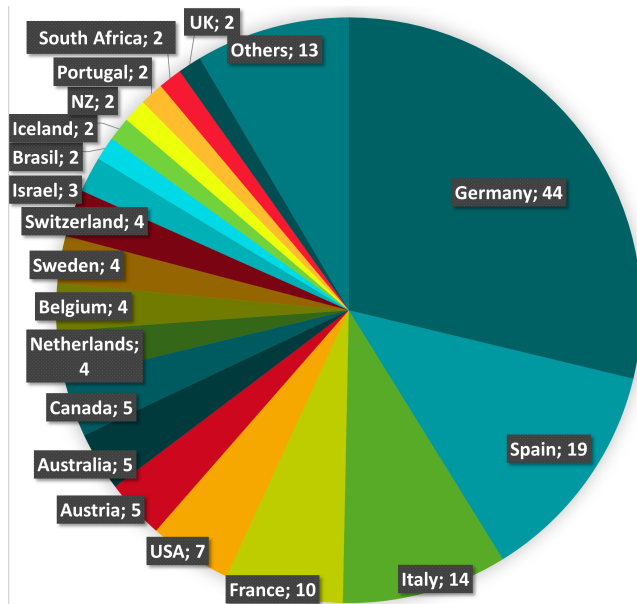


Fig. 3 Countries of survey participants

distinguished individuals, each of whom was instrumental in initiating the ER (Peter P.S. Chen), the MODELS (Jean Bézivin), and the BPM (Wil van der Aalst) conferences. The interviews focused on the objectives pursued at the time of the foundation of the conference and the vision of these outstanding scientists on the future of modeling and its communities.

4 Past and Present of Modeling: Main Results

This section is organized according to the analysis framework presented in Section 3 and includes the results of our survey and bibliometric analysis related to the past and present of modeling.

4.1 Foundations

To identify on which foundations the individual communities base their activities and to answer Q1.2, we provided free text fields in the survey in which the participants could indicate their up to ten most important fundamental works (scientific papers or books) (3.2 in Table 5). Table 2 shows the result. For the sake of clarity, we have limited ourselves to those works that were mentioned at least twice by one modeling community and have arranged the table horizontally according to the fundamental works of the individual communities, starting with a section in which works are listed that were mentioned by members of several communities. The columns indicate the individual number of mentions. It was somewhat surprising for us that only three

works reached or exceeded the threshold of 10 mentions, and that no work was mentioned more than three times by members of the MODELS community (although the cohort size of participants from the MODELS community was somewhat smaller).

4.2 Research Methodologies

To answer Q1.3, we asked the participants about which research methodologies they predominantly used (3.3 in Table 5) and to rank them according to their subjectively judged importance (3.4 in Table 5). For convenience, we pre-listed 17 common methodologies, although participants could also add additional ones that they felt were important. Eight participants made use of this option. Figure 4 shows the 17 suggested methodologies (x-axis) and by what percentage of the respective research community they were mentioned.

Overall, *Concept Implementation (proof of concept)* was the most frequently mentioned methodology, followed by *Design Science* and *Case Study*, with participants from the software modeling community placing greater emphasis on *Implementation* and *Case Studies* than on *Design Science*. For the BPM community, two more methodologies are highlighted: *Systematic Literature Review* (12,3%) and *Data Analysis* (13,1%). The latter plays only a minor role in the software modeling community (2,8%). All responses and the eight additionally proposed methodologies (each once) can be found in the online accompanying materials [10].

4.3 Research Topics

The subsequent analysis of research topics is separated into the results of the currently most exciting modeling research topics (Section 4.3.1) and those modeling research topics with a need for action (Section 4.3.2).

4.3.1 Currently Most Exciting Topics

To answer Q1.1 concerning the present, we asked in the survey *What are currently the most exciting topics and application areas in modeling for you?* (5.1 in Table 5). Since this was an open question, the answers were correspondingly heterogeneous. In the analysis, we therefore first had to carry out a coding of the terms given, e.g., concerning case sensitivity, synonyms, abbreviations, etc. We then filtered out the most frequently mentioned topics across all communities and entered them in descending order in Table 3, supplemented by the number of mentions in the individual communities. We restricted ourselves to those topics that had at least

Table 2 Foundational literature of the three modeling communities.

Literature	BPM	ER	MODELS
Foundational literature across the modeling research communities			
Chen, P.P.S. (1976). <i>The entity-relationship model – toward a unified view of data</i> [22]	1	16	2
Dumas, M., La Rosa, M., Mendling, J., & Reijers, H.A. (2013). <i>Fundamentals of business process management</i> [29]	12	1	-
Olivé, A. (2007). <i>Conceptual modeling of information systems</i> [60]	-	4	2
Stachowiak, H. (1973). <i>Allgemeine Modelltheorie</i> [65]	-	2	3
Brambilla, M., Cabot, J., & Wimmer, M. (2017). <i>Model-driven software engineering in practice</i> [12]	1	-	3
Foundational literature of the BPM community			
W. van der Aalst: <i>Process Mining – Data Science in Action</i> [75]	10	-	-
Mathias Weske, <i>BPM – Concepts, Languages, Architectures</i> [78]	6	-	-
Carmona, J., van Dongen, B., Solti, A., & Weidlich, M. (2018). <i>Conformance checking</i> [20]	2	-	-
Friedman, J. H. (2017). <i>The elements of statistical learning: Data mining, inference, and prediction</i> [34]	2	-	-
Foundational literature of the ER community			
Elmasri, R., & Navathe, S.B. (2000). <i>Fundamentals of Database Systems</i> [30]	-	6	-
Batini, C., Ceri, S., & Navathe S.B. (1992). <i>Conceptual database design: an entity-relationship approach</i> [6]	-	5	-
Codd, E.F. (2002). <i>A relational model of data for large shared data banks</i> [23]	-	3	-
Karagiannis, D., Mayr, H. C., & Mylopoulos, J. (2016). <i>Domain-specific conceptual modeling</i> [45]	-	3	-
Thalheim, B. (2010). <i>Entity-relationship modeling: foundations of database technology</i> [69]	-	3	-
Embley, D. W., & Thalheim, B. (Eds.). (2012). <i>Handbook of conceptual modeling: theory, practice, and research challenges</i> [31]	-	2	-
Ferstl, O. K., & Sinz, E. J. (2015). <i>Grundlagen der Wirtschaftsinformatik</i> [32]	-	2	-
Guarino, N. (1994). <i>The ontological level</i> [39]	-	2	-
Halpin, T., & Morgan, T. (2010). <i>Information modeling and relational databases</i> [41]	-	2	-
Kent, W. (1978). <i>Data and reality</i> [49]	-	2	-
Foundational literature of the MODELS community			
Broy, M., & Stølen, K. (2012). <i>Specification and development of interactive systems: focus on streams, interfaces, and refinement</i> [13]	-	-	2
Gamma, E., Johnson, R., Helm, R., Johnson, R. E., & Vlissides, J. (1995). <i>Design patterns: elements of reusable object-oriented software</i> [36]	-	-	2

three mentions overall. Thus, the table contains a total of 30 topics; that is about 30% of the topics mentioned and they combine 67% of all mentions. The total of 102 topics mentioned can be found in the accompanying materials [10].

Not surprising, *Artificial Intelligence (AI)*, *Conceptual Modeling* [58], and *Domain Specific Languages (DSLs)* [48] topped this list. It is interesting, however, that topics such as *Flexible Modeling* and *Model Integration*, which allow for working with different

tools and languages, were also mentioned relatively frequently.

Next, we were interested in whether certain topics are particularly prevalent in one or two modeling communities. Table 3 therefore shows not only the total number of mentions of a topic but also a breakdown with respect to the three communities. It can be derived, that some topics, such as *AI*, *DSL*, and *flexible modeling* seem to be considered relevant for all three communities whereas others only seem to play a role in one community. For example topics like *data model-*

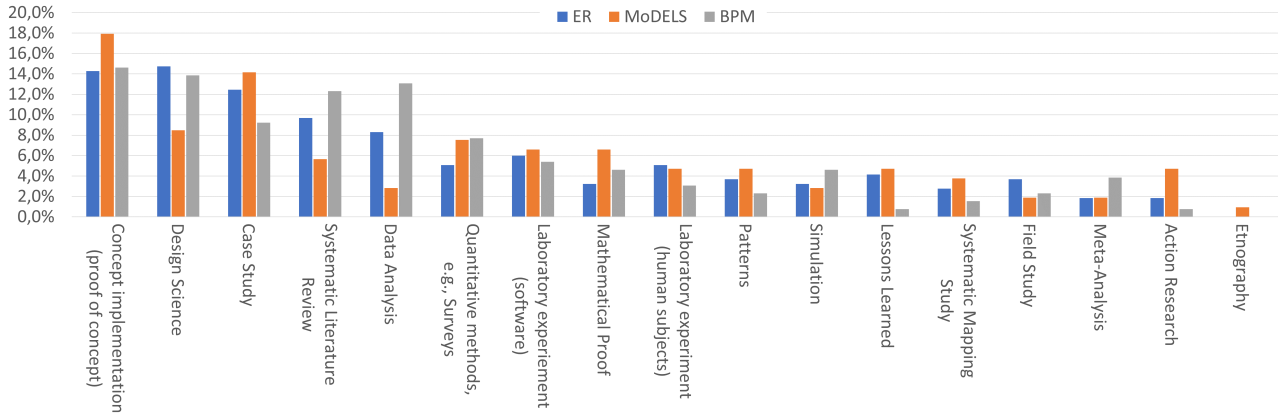


Fig. 4 Relative number of mentions of a research methodology, grouped by research community.

ing, *ontologies*, *NoSQL*, and *modeling theory* are prevalent in the ER community, whereas the BPM research community focuses on *process modeling*, *Industry 4.0*, and *simulation*. The MODELS community has prevalent topics with, e.g., *low code*, *language engineering*, and *testing*.

Of course, the absolute numbers of the respective mentions are not very high, but they nevertheless substantiate the assumption that each community has its special topics on the one hand, but on the other hand there are also many common topics.

The participants from the industry named a total of 24 topics that were of current interest to them. As with researchers, *AI* plays an important role (3 mentions), and *flexible modeling* (2 mentions) also appears to be significant for both groups. However, since the number of mentions on the part of the industry participants is rather low, a richer analysis is not possible. The remaining mentioned topics with the number of mentions are as follows: capability modeling (2), enterprise architecture (2), model composition (2), process modeling (2), big data (1), data analytics (1), data protection (1), DSL (1), formalization (1), Industry 4.0 (1), language engineering (1), model analysis (1), model integration (1), model processing (1), model quality (1), model verification (1), security (1), semantic web (1), state modeling (1), surrogate modeling (1), usability (1), and web modeling (1).

4.3.2 Topics with a Need for Action

This section focuses on those topics for which scientists and practitioners see a need for future action. We asked both respectively, “for which modeling topics do you see an explicit need for action?” (5.2 and 5.3 in Table 5).

A total of 189 topics [10] were mentioned by all participants. Table 4 lists those 22 topics that were mentioned at least three times by the researchers (from all communities); the topics mentioned by the practitioners are listed below. The data in Table 4 represents

Table 3 Most exciting topics by the researchers.

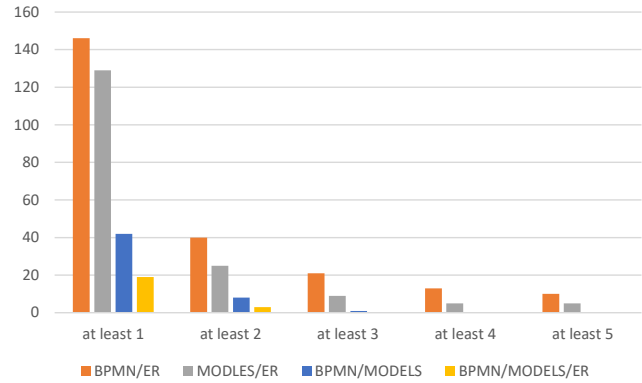
Topic	Mentions			
	Total	BPM	ER	MODELS
AI	21	5	11	5
conceptual modeling	12	1	10	1
DSL	10	2	3	5
data modeling	9	2	7	0
model-driven engineering	9	1	6	2
ontologies	9	0	9	0
process modeling	9	7	1	1
automation	7	4	1	2
flexible modeling	7	2	3	2
model integration	7	0	5	2
human factors	6	0	2	4
bioinformatics	5	0	5	0
low code	5	0	1	4
model transformations	5	0	3	2
modeling tools	5	0	3	2
NoSQL	5	0	5	0
agility	4	1	2	1
big data	4	0	4	0
enterprise architecture	4	1	3	0
knowledge graphs	4	0	4	0
language engineering	4	0	1	3
modeling theory	4	0	4	0
security	4	0	2	2
usability	4	0	2	2
Industry 4.0	3	3	0	0
model processing	3	1	2	0
modeling education	3	2	1	0
multi-level modeling	3	0	1	2
simulation	3	3	0	0
testing	3	0	0	3

Table 4 Most mentioned topics with a need for action by the researchers.

Topic	Mentions			
	Total	BPM	ER	MODELS
modeling tools	17	0	3	14
AI	10	1	7	2
human aspects	8	0	2	6
data modeling	7	1	6	0
industry application	7	0	5	2
process modeling	6	5	0	1
modeling in the humanities	5	0	5	0
usability	5	1	3	1
empirical research	4	0	2	2
interoperability	4	3	1	0
model execution	4	1	2	1
model integration	4	0	1	3
model-driven engineering	4	0	3	1
modeling education	4	0	2	2
NoSQL	4	0	4	0
automation	3	0	2	1
conceptual modeling	3	0	3	0
IoT	3	2	1	0
meta modeling	3	1	2	0
modeling theory	3	0	3	0
scalability	3	0	1	2
semantics	3	1	2	0

60.3% of all mentions and 26.5% of all topics—a complete list of all responses is available online [10]. Many similarities can be seen between the topics having an explicit need for action (Table 4) and the currently most exciting topics (Table 3), e.g., *AI*, *data modeling*, *process modeling*, *model integration*, and *human factors*. However, some discrepancies between the assessment of currently important topics and those that will be important in the future also appear to be noteworthy. For example, *domain-specific languages* and *ontologies* were very frequently mentioned as a current topic, while hardly any need for future action was seen. In contrast, *modeling tools* and *usability* were mentioned much more frequently as topics needing action compared to the mentions as current topics. A tentative conclusion, then, might be, that action is needed to focus on human factors, modeling tools, and their usability to achieve immediate positive impact.

Due to the low number of responses from industry participants, again no robust statements can be made. Except *model partitioning* all of the following topics

**Fig. 5** Numbers of authors with at least n publications in two or three community platforms

were mentioned once: *AI*, *automation*, *conceptual modeling*, *data modeling*, *debugging*, *derivative networks*, *DevOps*, *DSL*, *information modeling*, *language composition*, *legacy system*, *model transformation*, *model versioning*, *model-driven engineering*, *modeling education*, *modeling of law*, *ontologies*, *process modeling*, *reverse engineering*, *scalability*, *software comprehension*, and *web modeling*.

4.4 Contact between modeling communities

This section is premised on the assumption that one indicator of exchange between communities and cross-community collaboration is that authors publish not only in one (i.e., a core) communication platform. To answer Q2.1 such publications were identified in the course of our ad-hoc bibliographic analysis. We identified a total of 5,079 authors who have published at least once in the main proceedings of the ER, MODLES, and BPM conferences—in each case since their beginning until 2022. Whereas 1,364 of these authors have more than one publication, only 20 have published at least one publication in all three conferences. 257 authors have published in two of the three conferences, in particular 42 in BPM and MODELS (no author has more than three publications in both, one author has exactly one publication in both), 146 in BPM and ER (21 having more than three publications in both), and 129 in MODELS and ER (nine having more than three publications in both). Figure 5 shows the number of authors who have at least n (1..5) publications in two or all three platforms. The complete bibliometric analysis results are provided online [10].

Overall, it can be stated that authors with multiple publications have a “home conference” where they publish most of their papers, with many (1,087 in total) being represented only at this conference.

Additionally to the bibliometric analysis, we asked the participants if they perceive the three modeling communities as being closely connected (7.1 in Table 5), and if they should be more closely connected (7.2 in Table 5). A five-item Likert scale was provided for answering these questions. Figure 6 visualizes the responses on a 100% scale per research community, i.e., we normalized the relative ratio to enable comparability albeit the differences in the number of responses from each community. The figure shows, that more than 40% of the participants from all three communities consider them as being closely connected. Nevertheless around 80% of all participants of all three modeling communities think that the communities should be more closely connected. The complete agreement was strongest reported by the MODELS and lowest by the ER community. When considering the sum of the agreement statements, no significant differences can be seen among all three communities.

To answer Q2.3 we have further asked in our survey why or why not the participants think the communities should be more connected (7.3 in Table 5). This was an open question for which we harmonized and clustered the answers. The most common arguments for a closer connection were that the communities

- “operate on a common basis and software in reality” (18 answers),
- “should integrate (harmonize) models, languages, techniques, and tools” (17 answers),
- “cover different perspectives” (15 answers),
- “should use synergies” (12 answers),
- “have the same research topic and problems” (7 answers),
- “save effort and be efficient” (7 answers),
- “should cooperate to achieve goals and solve grand challenges” (6 answers),
- “are integrated in engineering process”, can “benefit from strengths”, and for a “better collaboration and co-creation” (all 5 answers).

Some participants highlighted negative aspects of the current situation by speaking “duplication”, scientists that work in “silos”, and “reinventing the wheel”. Arguments against a closer connection included:

- “methods and approaches are different” (5 answers),
- “members of the communities can work together anyway” (2 answers).

4.5 Cooperation between Research and Industry

Q2.4 refers to the contact and cooperation between practice and research. To answer it, we again used a

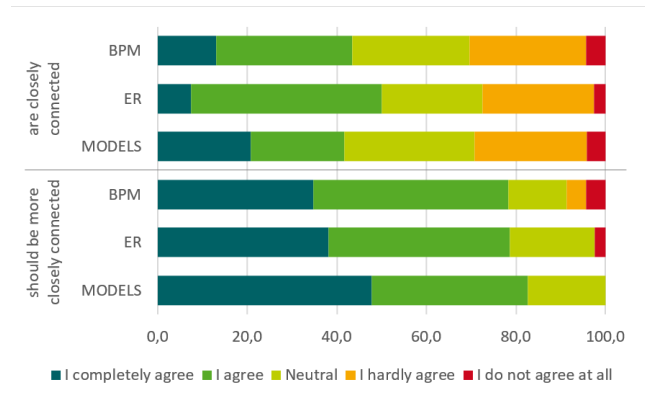


Fig. 6 Connections between the communities

combination of specific bibliometric analysis and three questions in our survey.

Regarding the bibliometric analysis, we limited ourselves to looking at the research papers—identified for answering Q2.1—of the main proceedings of the BPM, ER, and MODELS conferences in the last five years (2018-2022) to identify more recent collaborations by evaluating the author’s affiliations. In this context, we consider as results of true collaborations those papers that have at least one author from industry and one from academia. Please also note that the following analysis concentrates on the academia/industry collaboration. Consequently, papers that are entirely co-authored by authors with an industrial affiliation are not considered although also present at the considered conferences.

Figure 7 shows, that the total number of industry/research collaboration papers for the last five years is the highest for the MODELS conference (approx. nine such papers per year), followed by the ER and the BPM conference (approx. four and three such papers per year, respectively). In relative terms, over the period of five years, MODELS featured 48 out of 173 industry/research collaboration papers (28%), ER 20 out of 193 (10%), and BPM 14 out of 122 (11%).

In the survey, we attempted to address the question of whether modeling research and industry should collaborate more, what researchers expect from collaborating with practitioners, and conversely, what practitioners expect from researchers (8.1, 8.2, and 8.3 in Table 5, respectively). To the first question we received a total of 84 usable responses, more precisely: 71 from researchers (84.5%) and 13 from industry participants (14.5%). A large majority of the researchers (68 out of 71) emphasized the benefits of increased cooperation with one of the negators simply stating, that enough cooperation is ongoing already—thus not really taking a negative standpoint. Likewise, 12 out of the 13

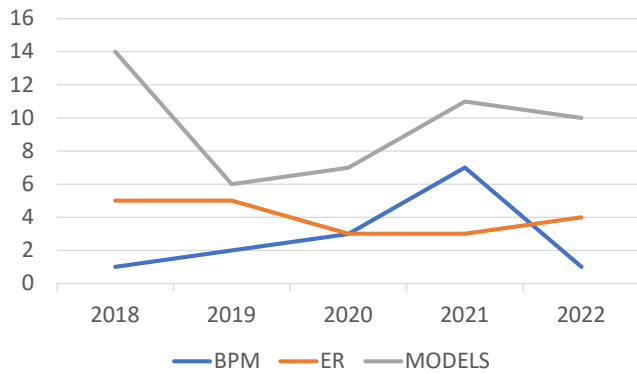


Fig. 7 Joint papers by at least one industry and one academic researcher in the three main conferences in 2018-2022.

industry participants voted for increased cooperation. The negator emphasized that cooperation not always makes sense because some of the industry problems “are boring from an academic perspective”. The two actual responses not supporting tighter cooperation between modeling research and industry stress that “Industrial Research is often not up-to-date” and that “industry strives only for short-term results”. Overall, both sides heavily emphasize the importance of cooperation.

Analyzing the responses about the expectations also yielded homogeneous results. Both, researchers and industry participants emphasized the importance of working on real problems and realistic cases (25 mentions from researchers and seven from the industry) and ensuring that modeling research remains relevant (21 mentions from researchers and four from the industry). Researchers generally emphasized the importance of knowledge transfer (19 mentions), i.e., realizing a “broader adoption”, “involving all stakeholders”, and “improving the communication with more practical fields”. From a research perspective, industrial collaboration shall also aid in applying the developed modeling languages and tools (14 mentions) and in evaluating proposed solutions in realistic settings (13 mentions).

When further decomposing the analysis to the three modeling communities, only a few differences have been observed. For example, working on real problems and realistic cases seemed equally expected from BPM researchers and MODELS researchers with 50% and 47.6% mentions, respectively, while only 22.2% of the ER researchers mentioned this in their expectation. The second interesting difference related to the knowledge transfer which was an expectation mentioned by 33.3% of the ER, 28.5% of the MODELS, and only 7.1% of the BPM researchers.

Taking a closer look at the responses from practitioners, they particularly emphasized better communi-

cation by means of industry-oriented communication of modeling research, e.g., “Simpler, shorter and more accessible descriptions of their work rather than just long technical papers written for other researchers” and a better alignment of the research topics to the needs and challenges of the industry, e.g., “More concern for current industry challenges rather than theoretical ideals”. Moreover, the practitioners expect improved automation, more efficient tooling with easier access, and interoperability between tools, e.g., “More interoperable, more accessible (JSON, web-ready, Typescript vs all in Java EMF); no need to install Eclipse”. Finally, practitioners are also seeking a tighter integration of modeling and model-driven engineering in (agile) processes, e.g., “Deeper connection to the practice, e.g., discussing the question of why agile processes don’t use modeling”.

5 Future and Vision of Modeling: Main Results

We addressed Q1.5 (“What does modeling have to achieve to be more important in 10 years”) in our survey with questions 6.1 and 6.2 (see Table 5). Both questions were to be answered in free text fields.

5.1 Modeling Future

For survey question 6.1, we received in total 74 responses, 11 from industry and 63 from academia. For a better overview, these have been grouped into the following 10 categories:

1. **who models**—responses that relate to specific user groups, e.g., modeling should be *used by practitioners, different domains, achieve a higher degree of interdisciplinary collaboration, or target different developer types (low, expert, no code)*;
2. **what is modeled**, e.g., *wide range of systems*;
3. **why modeling**, e.g., *handle complexity of reality, support communication and collaboration, support digital transformation, or to be human-relevant and that we can explicate benefits*;
4. **the modeling process**, e.g., *modeling should be integrated into the development lifecycle, or integrated with programming*;
5. **research methods**, e.g., *empirical research*;
6. **improvements on the level of (modeling) languages and DSLs**, e.g., *to create and use DSLs at all, to make DSL engineering easier and faster, or to provide foundations and methodologies*;
7. **improvements on the level of models**, e.g., *to enable and provide more automation, allow for (faster) generation/transformation from models,*

- and *model execution or simulation*, provide *reusable modeling components/model repository*, and *version management for models*;
8. **how to improve the use of modeling languages and models**, e.g., *provide tooling at all*, and provide *better tools (easy, no accidental complexity, good user experience)*;
 9. **education**, e.g., *better modeling education*;
 10. **connection to other areas of informatics**, e.g., *show the connection to AI or modeling for big data processing*.

The total of 11 participants from **industry** mentioned categories 7 and 8 most frequently. Four of them want to enable *more automation*, and three are looking for *better tools that are easy for end users to use, remove unintended complexity and provide a good user experience*. Other aspects, mentioned at least twice were related to category 4 (*modeling should be integrated into programming*) and 3 (*need for faster generation and transformation methods and tools*) and for *version management for models*). Higher-level categories 9 and 10 were not mentioned by practitioners.

Many responses of the **research** participants cover the categories 1, 3, 4, 7, and 8, with the top three responses (19 of 63) being *modeling should be used by practitioners, modeling should be human-relevant and benefits should be explicated*, and *better tools are needed*. Eight participants think that *modeling should be easier and become more usable*. Seven participants note that *we need to enable and increase the degree of automation*. Six out of 63 answers mentioned the need for (faster) generation and transformation from models, to improve model execution and simulation, and to provide tooling at all.

The 63 research responses are distributed among the communities as follows: ER 32, BPM 11, and MODELS 20. The following aspects were important to representatives of all three communities:

- modeling should be easier and become more usable,
- modeling should be used by practitioners (with more mentions from the ER and MODELS community),
- enable and provide more automation.

The responses from the ER community revealed a higher interest in *showing benefits for other disciplines*. BPM representatives see more importance in *integrating modeling into the development life-cycle*. MODELS representatives see an importance for *better tools and the need to provide tooling at all*. ER and BPM representatives mention in common that *modeling should be human-relevant* and that *we need to explicate benefits*. BPM and MODELS representatives mention

in common that *modeling should be integrated with programming*, the *need for DSLs*, (*faster*) *generation and transformation from models*, and the *need for model execution and simulation*.

5.2 Modeling Vision

Survey question 6.2 (“Describe your vision for modeling”) received 57 meaningful, non-empty responses, 10 from industry and 47 from research (6/26/15 from the BPM/ER/MODELS community, respectively).

Again, we clustered the individual responses, resulting in the following higher-level categories that

1. **relate to conceptual modeling itself**, e.g., *Raise the level of abstraction*;
2. **focus on what should be modeled**, e.g., *to model the human system, including cognition*;
3. **address the stakeholder and the domain to be modeled**, e.g., models as enablers for *communication in heterogeneous groups of stakeholders about complex (organizational) problems* and *modeling AI systems*;
4. **propose concrete steps to be performed to realize the vision**, e.g., realizing a *global platform for collaborative and open-source modeling*;
5. **address the value conceptual modeling and models may provide**, e.g., modeling should *support the full life-cycle of software and all activities involved, including requirements capture, design sketches, interactive model execution/animation, formal verification, code generation, reverse engineering, DevOps, etc*;
6. **address the assumptions underlying the vision**, e.g., the availability of processes and vast amounts of data.

Within these higher-level categories, further sub-categories have been defined subsequently while analyzing the responses. Interested readers can find this in the accompanying material [10].

The responses from the **industry representatives** were quite heterogeneous, covering all six categories. A focus was placed on category 4, especially on the question of *how modeling languages and supporting tools need to be further enhanced*. Four responses were related to *improving the accessibility and usability of modeling languages and tools*. Practitioners are calling for *tools that enable collaborative and concurrent modeling via the browser*. Another vision concerns *greater flexibility in the use of modeling languages*, for example, by *composing them and enabling reuse of models through openly available repositories*.

Many responses of the **research participants** relate to category 1, emphasizing the *need for a better understanding of the underlying theoretical and conceptual foundations* of conceptual modeling, category 2 *applying modeling in real-world situations*, category 3 *better alignment of the modeling languages to the modeled domain*, and category 4 *improving modeling education* to establish a wider awareness and acceptance of modeling by further stakeholders.

Adapting modeling languages to better cope with domain-specific aspects was mentioned in 14 out of the 47 responses. 10 researchers addressed the role of *models as a means of enabling communication amongst different stakeholders*. Similarly to the industry responses, *improving the accessibility and the ease of use of modeling languages and modeling tools, specifically web modeling tools* was addressed (category 4). In addition, 17 research participants foresaw the development of *new concepts for domain-specific applications*, and 11 addressed the *use of conceptual models to generate code*.

Despite the limited number of responses, some interesting differences in the answers from the specific communities could be filtered out. For example, *code generation* was mentioned by nine MODELS representatives (out of the overall 11 nominations in the survey). The ER representatives primarily (11 out of 26) expressed the vision of the *development of new concepts* representing 11 of the overall 17 mentions in the survey. The BPM representatives focused slightly more on *accessibility and usability of modeling languages and tools*, as well as on the *use of models as an enabler for communication amongst different stakeholders*. Finally, ER and MODELS representatives together expect increasing *applications of modeling in different domains*.

6 Expert Interviews

In each of the three communities we consider here, there are a number of outstanding representatives whose names are intuitively associated with the main conference in question because they were instrumental in its founding and development. To round off this paper, and to incorporate the knowledge and assessments of such personalities, we asked three of them for an interview: Wil van der Aalst for the BPM community, Peter P.S. Chen for the ER community, and Jean Bézivin for the MODELS community. In one case the interview was conducted in person, in the other two cases we submitted our questions in writing and received written responses. The questions were largely consistent, while in some cases we also tried to address community-specific aspects.

We received very extensive answers from these renowned personalities, the complete reproduction of which would go beyond the scope of this paper. Therefore, we have taken the liberty to filter out and summarize certain aspects that support or complement what has been said so far. The order of the transcriptions corresponds to the sequence of the foundation of the conferences concerned.

6.1 Peter P.S. Chen: the ER community

Peter P.S. Chen, as the “father” of the entity-relationship model, is of course one of the most important drivers of the community also known as ER. With his paper “The Entity-Relationship Model: Toward a Unified View of Data” [22] he intervened in the then-current discussion of data modeling paradigms to contribute to their harmonization. This is related in particular to the *CODASYL* Network model [68] and the relational model (cf. [25]). His work has been exceedingly successful and influential; as of June 2023, it has more than 13,250 citations, according to Google Scholar.

The inauguration of the ER community

In response to the publication of the initial paper, “*there was a lot of interest in extending or applications of the Entity-Relationship Model.*” It was therefore decided to establish a forum where a small group of interested researchers and practitioners would meet to exchange ideas and discuss ER-related challenges. This was the inauguration of the ER community with its first ER conference taking place in Los Angeles, California in 1979. Based on the surprisingly high attendance and the positive feedback from the first conference, it was decided to organize a second ER conference two years later in Washington. After the first four ER conferences which were all held bi-annually in the US, the conference moved to an annual scheme with changing international hosts.

From Practice-oriented to Theory-focused

While the ER conference was initially aimed to bring together researchers and practitioners, the focus shifted toward the theoretical foundations of conceptual modeling in the following years. Peter Chen is sure that this is also one of the reasons why the research interest in the US declined to some extent while the interest in the EU significantly increased. He consequently emphasizes the need to rebuild strong relationships between research and practice: “*I think the modeling research should seek closer collaboration with industry/application. The researchers can expect to get a better understanding of i) the real issues the*

practitioners are concerned about, and ii) the obstacles in testing the theories or proposed techniques.” Likewise, he also emphasizes the need to establish inter-community collaboration across the different modeling communities by e.g., having “both, common forums (platforms) and separate forums co-exist and serve different purposes.”

Maintaining & increasing the relevance of modeling

Peter strongly disagrees with the claim that “modeling is out” and expects modeling to “have a bright future.” Based on his expertise, and again emphasizing his previous arguments, “the community needs to i) seek a closer collaboration with industry/practitioners, and ii) increase cross-fertilization with other communities.”

Current hot topics

Peter believes that amongst the many open and important questions, the question of “how to integrate data, software, and processes” is of central interest to the ER community. As a natural continuation of his impactful paper on the Entity Relationship model, Peter describes this research direction with the aim to develop a “unified view of data, software, and processes.”

Advice for PhD students

When asked what he would recommend to young PhD students, Peter replied: “You have chosen a very interesting and exciting field to work on. Work hard, be patient, and you will have a bright future. Remember: Rome was not built in one day!” Some concrete advice was also provided: “Split your time in two parts. For a part of your time, you will work on the hot topics suggested by others; for the other part of your time, you will work on the topic you think are important.”

6.2 Jean Bézivin: the MODELS community

Jean Bézivin is one of the founders of the initially called UML and now MODELS conference. His most influential papers have more than 1,400 citations. He started his career as an Assistant Professor in Rennes, where he got involved in object-oriented programming and the OOPSLA conference. To capture the correspondence between real-world objects and programming objects more than 50 OO-modeling languages arrived in a short period of time. For industry, this was a very unsatisfactory situation. Therefore, OMG started a unification process: the birth of UML. Jean remembers a hot discussion about the “frustration of academic researchers not being able to influence the OMG decisions” on important modeling issues. Thus, they needed their own place to discuss—an independent academic conference.

The emergence of the MODELS conference

This endeavor started as “UML’98 International Work-

shop with the support of OMG” in Mulhouse (in the midst of a huge French strike) and one year later as a UML conference series one year in Europe and one year outside Europe⁷. Contrary to initial expectations there was less cooperation than foreseen between the academic and industrial events, and the concern came up that the conference’s development was highly connected with the relevance of UML. “We realized that the future of modeling was considerably larger than the future of UML” and renamed the conference series from 2005 on to “Model Driven Engineering of Languages and Systems” (MODELS), preferred to Model Engineering of Software and Systems which had a disputable acronym”.

Relationship between research and practice

Only constant contact between research and industry enables university researchers to solve real problems. Jean, therefore, experiences the interplay between small academic research groups, big industrial players that wanted to investigate the applicability of modeling ideas on their applications, normative organizations, and open-source communities as rather complex, challenging, and sometimes resource-consuming but it “created interesting interactions and highly positive results”.

Current hot topics

Jean first responded to the question about his opinion of hot topics with the counter-question of what we meant by this, “e.g., a topic on which a lot of money will be available for projects in the five coming years?” A more consensual answer, however, would probably be topics such as requirements engineering, systems engineering, artificial intelligence, machine learning, cybernetics, or cybersecurity, “but one should not forget that hot topics of today may become cold topics soon”.

Unification property and interdisciplinarity

In Jean’s opinion, the essential core property of model engineering is unification “i.e., the possibility to capture a lot of different phenomenon or situations within the same regular framework”. Therefore, “the duality of modeling processes and products should be studied more deeply. This requires to model aspects individually but also to provide bridges between these different perspectives and aspects.” Up to now, there exists a huge variety of modeling languages in different disciplines as they help us to understand the world, and provide us with the opportunity to use software modeling as a support for interdisciplinarity. In future, “modeling (in a broad meaning) could be taught at middle schools, as a fundamental discipline like Mathematics, Physics, or Geography. When interdisciplinarity will become a key subject,

⁷ <https://dblp.uni-trier.de/db/conf/models/index.html>

i.e., a first-class discipline on itself, model engineering will find the place it deserves."

Advice for PhD students

Reading a lot and being innovative is Jeans' main advice for PhD students: *"We are not yet at a point where the research is going to be incremental, more likely it will be in rupture, so dare to be iconoclast. Look at what has been done in other disciplines. Dig into the old pre-UML literature to find some good ideas buried there."*

6.3 Wil van der Aalst: the BPM community

Wil van der Aalst is one of the founders of the BPM conference. According to ORCID as of June 2023, he has contributed to 1,462 publications to date, which have been cited a total of more than 135,000 times. At the beginning of his career, he was concerned with simulation tools for the specification and simulation of software systems. Over time it turned out that these techniques were more suited for business processes or workflow processes than for the description of software. Thus, he evolved into Petri Net modeling. As he says about himself *"I'm a Petri Net person. Even my children are drawing Petri Nets."*

The emergence of the BPM community

As Wil found the practical application of Petri Net in workflow management technology increasingly important, he organized, together with Arthur ter Hofstede and Mathias Weske, the first BPM conference⁸ in 2003 in Eindhoven, the Netherlands, in conjunction with the 24th International Conference on Application and Theory of Petri Nets⁹. *"What was surprising is that already in the first year, it had a size approximately the same as the Petri Net conference itself. What was also very clear is that there was immediately an interest from industry, so several workflow vendors, etc., were there."* Since its launch, the BPM Conference has grown year by year, and as with the other conferences, workshops supplemented the program. While there are BPM researchers in the USA, Russia and China, but no structured community, the European BPM community grew continuously in various dimensions: *theoretical, systems, and a more managerial orientation.* With the *process mining* community, *data orientation* has been added. After some time, process-mining-related papers at the BPM conference went to more than half of it. *"One could see that as an unhealthy development because clearly, BPM is more than just process mining."* The decision was taken to have a separate conference, and in 2019,

the process mining conference (ICPM)¹⁰ was started in Aachen, Germany. Up to now, many researchers are in both the BPM and Process Mining communities.

Relationship between research and practice

The BPM and process mining community seem to have no problem with relevance in practice: *"What was inspiring is to be involved with all of these companies doing workflow projects that miserably failed for various reasons and the completely unprofessional way in which people would select these types of systems not realizing what kind of limitations those had."* This interaction was especially brought to life through the meetings of the workflow management coalition which described workflow system capabilities independent of a particular application domain. *"Process mining companies are growing like hell. They are very greedy to adopt ideas in systems because everybody can see a lot of business value in it."* And: *"One of the significant differences between process mining and traditional BPM is that you can only research if you have data. You are forced to do practically relevant things."* But research must be aware that this good connection to practice might be a temporary development. And it should be clear that research is interested in more generic challenges and the transferability of solutions: *"I think that the modeling community should be open for the needs of industry but not a specific need."*

Collaboration across community boundaries

The different modeling communities should share their ideas and interact. However, for progressing in a field, one has to work in a very mono-disciplinary way and focus on specific questions. This makes cooperation more difficult because colleagues from other communities do not understand the respective challenges. In the worst case, they then even reject papers from another community.

Current hot topics

Wil considers topics such as automation in a data-driven sense and object-centric process mining in particular to be challenging and promising for the future. This includes, e.g., robot process automation which has a mixture of being data-driven, fact-based, and focuses on new forms of automation. Visionary things in a 10-years perspective are, e.g., to support world-wide production labs (in the context of the RWTH Aachen Cluster of Excellence Internet of Production) and digital twins. Besides these practical areas also many foundational problems remain, or as he would formulate it: *"I've been working on the same problems for 20 years and they are still not solved."*

Advice for PhD students

"Don't follow the crowd, try to do something original."

⁸ <https://dblp.uni-trier.de/db/conf/bpm/index.html>

⁹ <https://dblp.uni-trier.de/db/conf/apn/index.html>

¹⁰ <https://dblp.org/db/conf/icpm/index.html>

At the same time, you should always be able to explain what you are doing in the real world.”

7 Discussion

The insights we gained from the survey, the bibliometric analysis, and the expert interviews showed that there are some agenda-setting topics the modeling community with its sub-divisions should consider in the future – and this would be best done together. This requires knowledge transfer (i.e., technologies, concepts, methods, and tools) between the modeling communities and to focus the cooperation on common interests. Based on the results presented at the outset, we sketch some of these common interests.

7.1 Modeling Tools

When talking about modeling tools, we have to take the differences in the communities into account. The **BPMN community** seems to be satisfied with their tooling in comparison to the MODELS and ER community. This might result from either a lower heterogeneity in the used languages or sufficient functionality provided by tools such as the Camunda Platform [80]. The **ER community** uses and develops various tools with a focus on graphical representations of Domain-Specific Languages (DSLs). This includes either proprietary developed tools, platforms enabling one to define own tools such as ADOxx [79], or DireWolf [51], or tools for ontology editing and visualization such as Protégé [82] and the OntoUML Lightweight Editor (OLED) [40]. Even though there is a higher variety of DSLs in this community, they are often representing data and data structures either with more problem space focus or more software solution focus. For the **MODELS community**, engineering tools are one of the core areas of research which is also reflected in how the main conference treats tool demonstrations. They are an integral part of the main program, presented in sessions together with technical research papers and journal-first publications.

There exist several commercial and research-driven language workbenches, e.g., grammar-based language workbenches such as MontiCore [42], Neverlang [74], Rascal [67], Spoofox [46], and Xtext [7], metamodel-based language workbenches such as EMF [66], GEMOC Studio [26], and MetaEdit+ [73], language workbenches which create the Abstract Syntax Tree (AST) directly such as JetBrains MPS [81], and a new breed of language workbenches specialized to develop Web modeling tools such as the Eclipse Graphical

Language Server Platform (GLSP) [5] which enable highly flexible modeling editors [9] with rich graphical user interfaces due to the Web technology stack they use [19, 56].

This heterogeneity of DSLs and tools leads to two challenges, which were already identified and stressed in 2007 [33]. First, the *enhanced tooling challenge* still occurs as, in addition to editors and code generators which are nowadays provided out-of-the-box for DSLs, further tooling such as advanced analyzers, debuggers, and testing tools are required which have to co-evolve along with the evolution of the DSLs. Second, the *DSLs-Babel challenge* concerns the usage of several DSLs in combination which is nowadays even becoming a larger problem as more and more DSLs are being built. Thus, we are still facing interoperability, language versioning, and language migration issues which require dedicated solutions—especially when it comes to industrial adoption. Facing these challenges as communities together and exchanging knowledge would help to improve the modeling research in our fields and beyond.

7.2 Modeling and AI

AI was often mentioned as a current topic in the survey and recent developments such as the new version of ChatGPT have led to extensive discussions in the modeling community [18]. The interest manifests itself in workshops at each of the main conferences, e.g., the MODELS workshop on AI and MDE (MDEIntelligence) [3], the ER workshop conceptual modeling meets AI (CMAI) [1], and the BPM workshop AI for business process management [4]. Not only improving modeling methods in AI, but also using AI for improved engineering processes are challenging research areas. This impression is reinforced by recent articles [8]. A Communications of the ACM article [77] prophesies the end of programming because of AI and in recent discussions, [15] posed the question if Large Language Models will replace modelers and code generators. Since its beginning, Informatics is a science characterized by continuous and disruptive changes and several technological (r)evolutions. Consequently, the modeling communities must also adapt to, contribute to, and drive such changes. However, this also requires maintaining and promoting a heterogeneous research landscape as well as discussing the general role of modeling for society.

7.3 Modeling and Human Aspects

It is not an easy task to capture the complexity of the real-world and especially humans and their needs

in models and modeling. However, the diversity of our users and the diversity of modelers require that we take these differences into account when creating software systems. To the best of our knowledge, current modeling research has some contributions to this topic. For instance, the authors of [38] work on human-centric topics for MDE, the MODELS community runs a workshop on Human Factors in Modeling and Modeling Human Factors [2], and in the ER community, there exists work on modeling humans for behavior assistance [57]. If we have a look at research from other modeling communities, e.g., Kofod-Petersen and Cassens work on context modeling including the personal and social context of users [50], we may notice that their modeling approaches of human factors might be useful for our communities as well. Thus, cooperation in modeling topics not only across the modeling sub-divisions but also across disciplinary boundaries is an essential undertaking for the future, especially when it comes to human aspects.

7.4 Research and Industry Cooperation

There seems to be a huge interest on both ends, so we believe efforts should be made to strengthen the cooperation between modeling researchers and the industry. At universities, there is currently also a strong movement to strengthen the third mission activities, and the established technologies are mature enough to be tested in industrial settings. Thus, in our research we are able to tackle real-world problems to create meaningful contributions to society, and at the same time, performing basic research on modeling and associated technologies. However, we also need to provide different abstraction levels for problems and their solutions to be able to apply a solution also for similar problems in other contexts with stable technologies. Besides individual cooperations, all main conferences are strong in providing space for academic and industry exchange, e.g., by providing industry days, or by offering a practice track such as the one of the MODELS conferences focusing on contributions from and with industry. However, additional formats may be established which also allow to apply the research results from the modeling communities in other disciplines as there is currently a high need in mostly all areas such as smart manufacturing, energy, transport, construction, cities, etc.

8 Conclusion and Future Steps

In this paper, we shed some light on the three modeling sub-divisions, and we elaborate on their common-

alities and differences with respect to the foundations, topics, and methods. We further provide insights into the current trends and the visions for the future both topic-wise and also with respect to a potential increase of collaboration across modeling sub-divisions and between research and industry.

All insights derived in this paper are of course constrained by some limitations, and thus, many questions remain open for subsequent, in-depth research. First, we have limited our bibliometric analysis to the three main conferences. However, to include all relevant conferences and journals would have been beyond the scope of this paper, as would have been a content analysis of the papers in these conferences and journals. Furthermore, it would also be interesting, for example, to determine whether and how intensively there are movements between the communities over time. Analysis attempts in that direction, but not focused on the modeling communities, have been proposed already [16]. Another limitation relates to the number of participants in our survey and their distribution across the modeling sub-divisions as well as between research and practice. Future research should extend the data basis, thereby challenging and/or updating our findings.

We hope that this paper gives an impetus to further study the three modeling sub-divisions of data, process, and software modeling. A major question now is how these modeling communities could strengthen each other to better address the challenges ahead and realize the visions for a prospective modeling future. For increasing the exchange of ideas, concepts, techniques, and technologies, further meetings may be targeted such as dedicated workshops to discuss cross-community applications, Dagstuhl seminars to identify the grand challenges in modeling, etc. Pragmatic formats may be, e.g., rotating workshops, summer schools involving speakers from different communities, invited keynotes, and cross-community research networks, seminars, and projects. Moreover, special issues in journals that explicitly invite contributions with perspectives from different communities may be another direction.

Having further discussions on such ideas, however, is of major importance for the further development of the modeling research community overall. This is a call for dedicated community efforts such as building meta-committees of representatives from the sub-communities of modeling, providing an open forum to further discuss and develop ideas, and identifying common interests such as cross-cutting emerging research topics which impact all sub-communities, e.g., digital twins to just name one prominent example.

Acknowledgements We thank all the participants in the survey as well as Wil van der Aalst, Peter P.S. Chen, and Jean Bézivin for their valuable contributions, without which this paper would not have been possible. We further want to thank the Leibniz Center for Informatics (Wadern, Germany) for accepting our Dagstuhl Seminar 18471 on “Next Generation Domain Specific Conceptual Modeling: Principles and Methods” and all participants of that seminar for initial discussions on this topic.

References

- Conceptual Modeling Meets Artificial Intelligence (CMAI). <https://workshop-cmai.github.io/2021/> (2021). [Online; accessed 06-August-2023]
- Human Factors in Modeling / Modeling of Human Factors (HuFaMo). <https://research.webs.upv.es/hufamo22> (2022). [Online; accessed 06-August-2023]
- Workshop on Artificial Intelligence and Model-driven Engineering (MDEIntelligence). <https://mde-intelligence.github.io> (2022). [Online; accessed 06-August-2023]
- AI for Business Process Management (AI4BPM). <https://sites.google.com/unitn.it/ai4bpm-2023> (2023). [Online; accessed 06-August-2023]
- The Eclipse Graphical Language Server Platform (GLSP). <https://www.eclipse.org/glsp> (2023). [Online; accessed 06-August-2023]
- Batini, C., Ceri, S., Navathe, S.B.: *Conceptual Database Design: An Entity-Relationship Approach*. Benjamin/Cummings (1992)
- Bettini, L.: *Implementing Domain-Specific Languages with Xtext and Xtend*. Packt Publishing Ltd (2016)
- Bork, D., Ali, S.J., Roelens, B.: Conceptual modeling and artificial intelligence: A systematic mapping study. *CoRR abs/2303.06758* (2023). URL <https://doi.org/10.48550/arXiv.2303.06758>
- Bork, D., Langer, P., Ortmayr, T.: A vision for flexible GLSP-based web modeling tools. *CoRR* (2023). URL <https://arxiv.org/submit/4987271/view>
- Bork, D., Michael, J., Wimmer, M., Mayr, H.C.: *borkdominik/ModelingCommunities: Online supplementary material repository* (2023). URL <https://github.com/borkdominik/ModelingCommunities>
- Bourque, P., Dupuis, R., Abran, A., Moore, J.W., Tripp, L.L.: The guide to the software engineering body of knowledge. *IEEE Softw.* **16**(6), 35–44 (1999). DOI 10.1109/52.805471
- Brambilla, M., Cabot, J., Wimmer, M.: *Model-Driven Software Engineering in Practice, Second Edition. Synthesis Lectures on Software Engineering*. Morgan & Claypool Publishers (2017). DOI 10.2200/S00751ED2V01Y201701SWE004
- Broy, M., Stølen, K.: *Specification and Development of Interactive Systems – Focus on Streams, Interfaces, and Refinement*. Monographs in Computer Science. Springer (2001). URL <https://doi.org/10.1007/978-1-4613-0091-5>
- Burgueño, L., Ciccozzi, F., Famelis, M., Kappel, G., Lambers, L., Mosser, S., Paige, R.F., Pierantonio, A., Rensink, A., Salay, R., Taentzer, G., Vallecillo, A., Wimmer, M.: Contents for a model-based software engineering body of knowledge. *Software and Systems Modeling* **18**(6), 3193–3205 (2019). DOI 10.1007/s10270-019-00746-9
- Cabot, J.: From models-to-code to models-to-prompts: the next MDE revolution? <https://modeling-languages.com/models-to-code-models-to-prompts/> (2022)
- Cabot, J., Izquierdo, J.L.C., Cosentino, V.: Are CS conferences (too) closed communities? *Commun. ACM* **61**(10), 32–34 (2018). DOI 10.1145/3209580
- Cabot, J., Vallecillo, A.: Modeling should be an independent scientific discipline. *Software and Systems Modeling* **21** (2022). DOI 10.1007/s10270-022-01035-8
- Cámara, J., Troya, J., Burgueño, L., Vallecillo, A.: On the assessment of generative AI in modeling tasks: an experience report with ChatGPT and UML. *Software and Systems Modeling* **22**(3), 781–793 (2023). DOI 10.1007/s10270-023-01105-5
- Carlo, G.D., Langer, P., Bork, D.: Advanced visualization and interaction in GLSP-based web modeling: realizing semantic zoom and off-screen elements. In: *Proceedings of the 25th International Conference on Model Driven Engineering Languages and Systems (MODELS)*, pp. 221–231. ACM (2022). DOI 10.1145/3550355.3552412
- Carmona, J., van Dongen, B., Solti, A., Weidlich, M.: *Conformance checking*. Springer (2018)
- Chakraborty, T., Dalmia, A., Mukherjee, A., Ganguly, N.: Metrics for community analysis: A survey. *ACM Comput. Surv.* **50**(4), 54:1–54:37 (2017). DOI 10.1145/3091106
- Chen, P.P.S.: The entity-relationship model—toward a unified view of data. *ACM transactions on database systems (TODS)* **1**(1), 9–36 (1976)
- Codd, E.F.: A relational model of data for large shared data banks. In: *Software pioneers*, pp. 263–294. Springer (2002)
- Colanzi, T.E., Vergilio, S.R., Assunção, W.K.G., Pozo, A.T.R.: Search Based Software Engineering: Review and analysis of the field in Brazil. *J. Syst. Softw.* **86**(4), 970–984 (2013). DOI 10.1016/j.jss.2012.07.041
- Date, C.J., Codd, E.F.: The relational and network approaches: Comparison of the application programming interfaces. In: G. Altshuler, R. Rustin, B.D. Plagman (eds.) *Proceedings of 1974 ACM-SIGMOD Workshop on Data Description, Access and Control*, pp. 83–113. ACM (1974). DOI 10.1145/800297.811532
- Degueule, T., Combemale, B., Blouin, A., Barais, O., Jézéquel, J.M.: Melange: A Meta-language for Modular and Reusable Development of DSLs. In: *Proceedings of the 8th International Conference on Software Language Engineering (SLE)* (2015)
- Delcambre, L.M.L., Liddle, S.W., Pastor, O., Storey, V.C.: A reference framework for conceptual modeling. In: *Proceedings of the International Conference on Conceptual Modeling (ER)*, pp. 27–42. Springer (2018)
- Donthua, N., Kumar, S., Mukherjee, D., Pandey, N., Lim, W.M.: How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research* **133** (2021). DOI 10.1016/j.jbusres.2021.04.070
- Dumas, M., La Rosa, M., Mendling, J., Reijers, H.A., et al.: *Fundamentals of business process management, vol. 1*. Springer (2013)
- Elmasri, R., Navathe, S.B.: *Fundamentals of Database Systems, 3rd Edition*. Addison-Wesley-Longman (2000)
- Embley, D.W., Thalheim, B.: *Handbook of conceptual modeling: theory, practice, and research challenges*. Springer (2012)
- Ferstl, O.K., Sinz, E.J.: *Grundlagen der Wirtschaftsinformatik*. Oldenbourg Wissenschaftsverlag (2015)

33. France, R., Rumpe, B.: Model-driven Development of Complex Software: A Research Roadmap. *Future of Software Engineering (FOSE '07)* pp. 37–54 (2007)
34. Friedman, J.H.: *The elements of statistical learning: Data mining, inference, and prediction*. Springer open (2017)
35. Friedrich, T., Schlauderer, S., Weidinger, J., Raab, M.: On the research paradigms and research methods employed in the BISE journal - A ten-year update. In: *Towards Thought Leadership in Digital Transformation: 13. Internationale Tagung Wirtschaftsinformatik, WI 2017* (2017)
36. Gamma, E., Helm, R., Johnson, R.E., Vlissides, J.: *Design patterns: elements of reusable object-oriented software*. Addison-Wesley Professional (1995)
37. Garousi, V.: A bibliometric analysis of the turkish software engineering research community. *Scientometrics* **105**(1), 23–49 (2015). DOI 10.1007/s11192-015-1663-x
38. Grundy, J., Khalajzadeh, H., Mcintosh, J.: Towards Human-centric Model-driven Software Engineering. In: *15th Int. Conf. on Evaluation of Novel Approaches to Software Engineering (ENASE'20)*, pp. 229–238. INSTICC, SciTePress (2020). DOI 10.5220/0009806002290238
39. Guarino, N.: The ontological level. In: *Philosophy and the cognitive sciences*, pp. 443–456 (1994)
40. Guerson, J., Sales, T.P., Guizzardi, G., Almeida, J.P.A.: OntoUML Lightweight Editor: A Model-Based Environment to Build, Evaluate and Implement Reference Ontologies. In: *Proceedings of the 19th IEEE International Enterprise Distributed Object Computing Workshop, EDOC Workshops*, pp. 144–147. IEEE Computer Society (2015). DOI 10.1109/EDOCW.2015.17
41. Halpin, T., Morgan, T.: *Information modeling and relational databases*. Morgan Kaufmann (2010)
42. Hölldobler, K., Kautz, O., Rumpe, B.: *MontiCore Language Workbench and Library Handbook: Edition 2021*. Aachener Informatik-Berichte, Software Engineering, Band 48. Shaker Verlag (2021)
43. Hölldobler, K., Michael, J., Ringert, J.O., Rumpe, B., Wortmann, A.: Innovations in model-based software and systems engineering. *The Journal of Object Technology* **18**(1), 1–60 (2019). DOI 10.5381/jot.2019.18.1.r1
44. Kandjani, H., Bernus, P.: The enterprise architecture body of knowledge as an evolving discipline. In: J. Cordeiro, L.A. Maciaszek, J. Filipe (eds.) *Proceedings of the 14th International Conference on Enterprise Information Systems (ICEIS)*, pp. 452–470. Springer (2012). DOI 10.1007/978-3-642-40654-6_27
45. Karagiannis, D., Mayr, H.C., Mylopoulos, J. (eds.): *Domain-Specific Conceptual Modeling, Concepts, Methods and Tools*. Springer (2016). DOI 10.1007/978-3-319-39417-6
46. Kats, L.C., Visser, E.: The Spoox Language Workbench: Rules for Declarative Specification of Languages and IDEs. In: *25th Annual ACM SIGPLAN Conf. on Object-Oriented Programming, Systems, Languages, and Applications (OOPSLA 2010)*, pp. 444–463. ACM (2010). DOI 10.1145/1932682.1869497
47. Katsikeas, S., Johnson, P., Ekstedt, M., Lagerström, R.: Research communities in cyber security: A comprehensive literature review. *Comput. Sci. Rev.* **42**, 100,431 (2021). DOI 10.1016/j.cosrev.2021.100431
48. Kelly, S., Tolvanen, J.: *Domain-Specific Modeling - Enabling Full Code Generation*. Wiley (2008). URL <http://eu.wiley.com/WileyCDA/WileyTitle/productCd-0470036664.html>
49. Kent, W. (ed.): *Data and Reality*, 1st edition. North-Holland (1978)
50. Kofod-Petersen, A., Cassens, J.: Using activity theory to model context awareness. In: *Modeling and Retrieval of Context*, pp. 1–17. Springer (2006)
51. Koren, I., Klamma, R., Jarke, M.: *Direwolf Model Academy: An Extensible Collaborative Modeling Framework on the Web*. In: *Companion Proc. of Modellierung 2020*, vol. 2542, pp. 213–216. CEUR-WS.org (2020)
52. Krenn, M., Zeilinger, A.: Predicting research trends with semantic and neural networks with an application in quantum physics. *CoRR* **abs/1906.06843** (2019). URL <http://arxiv.org/abs/1906.06843>
53. Liu, X., Bollen, J., Nelson, M.L., de Sompel, H.V.: Co-authorship networks in the digital library research community. *Inf. Process. Manag.* **41**(6), 1462–1480 (2005). DOI 10.1016/j.ipm.2005.03.012
54. Mäntylä, M.V., Graziotin, D., Kuuttila, M.: The evolution of sentiment analysis - A review of research topics, venues, and top cited papers. *Comput. Sci. Rev.* **27**, 16–32 (2018). DOI 10.1016/j.cosrev.2017.10.002
55. Mayr, H.C., Thalheim, B.: The triptych of conceptual modeling. *Software and Systems Modeling* **20**(1), 7–24 (2021). DOI 10.1007/s10270-020-00836-z
56. Metin, H., Bork, D.: On Developing and Operating GLSP-based Web Modeling Tools: Lessons Learned from bigUML. In: *Proceedings of the 26th International Conference on Model Driven Engineering Languages and Systems (MODELS)*. IEEE (2023)
57. Michael, J., Mayr, H.C.: Conceptual modeling for ambient assistance. In: *Conceptual Modeling - ER 2013, LNCS*, vol. 8217, pp. 403–413. Springer (2013)
58. Mylopoulos, J.: *Conceptual modelling and Telos. Conceptual Modelling, Databases, and CASE: an Integrated View of Information System Development* pp. 49–68 (1992)
59. Nart, D.D., Degl’Innocenti, D., Pavan, A., Basaldella, M., Tasso, C.: Modelling the user modelling community (and other communities as well). In: *User Modeling, Adaptation and Personalization - 23rd International Conference, UMAP 2015*, pp. 357–363. Springer (2015). DOI 10.1007/978-3-319-20267-9_31
60. Olivé, A.: *Conceptual modeling of information systems*. Springer Science & Business Media (2007)
61. Otte, E., Rousseau, R.: Social network analysis: a powerful strategy, also for the information sciences. *Journal of Information Science* **28**(6) (2002). DOI 10.1177/016555150202800601
62. Robinson, S., Arbez, G., Birta, L.G., Tolk, A., Wagner, G.: Conceptual modeling: Definition, purpose and benefits. In: *Winter Simulation Conference (WSC)*, pp. 2812–2826 (2015)
63. Sciabolazza, V., Vacca, R., Okraku, T.K., McCarty, C.: Detecting and analyzing research communities in longitudinal scientific networks. *PLoS ONE* **12**(8) (2017). DOI 10.1371/journal.pone.0182516
64. Scott, J.: *Social network analysis: developments, advances, and prospects*. Springer Nature Social Network Analysis and Mining **1** (2011). DOI 10.1177/016555150202800601
65. Stachowiak, H.: *Allgemeine modelltheorie*. Springer (1973)
66. Steinberg, D., Budinsky, F., Merks, E., Paternostro, M.: *EMF: Eclipse Modeling Framework*. Addison-Wesley (2008)
67. van der Storm, T.: *The Rascal Language Workbench*. CWI. Software Engineering [SEN] (2011)

68. Taylor, R.W., Frank, R.L.: CODASYL data-base management systems. *ACM Comput. Surv.* **8**(1), 67–103 (1976). DOI 10.1145/356662.356666
69. Thalheim, B.: Entity-relationship modeling – foundations of database technology. Springer (2000)
70. Thalheim, B.: The conceptual model \equiv an adequate and faithful artifact enhanced by concepts. In: *Information Modelling and Knowledge Bases XXV*, vol. 260, pp. 241 – 254. IOS press (2014). DOI 10.3233/978-1-61499-361-2-241
71. Thalheim, B.: Conceptual model notions - A matter of controversy: Conceptual modelling and its lacunas. *Enterp. Model. Inf. Syst. Archit. Int. J. Concept. Model.* **13**(Special), 9–27 (2018). DOI 10.18417/emisa.si.hcm.1
72. Thalheim, B., Nissen, I.: *Wissenschaft und Kunst der Modellierung*. De Gruyter, Berlin, Boston (2015). URL <https://www.degruyter.com/view/title/505797>
73. Tolvanen, J.P., Kelly, S.: MetaEdit+: Defining and Using Integrated Domain-Specific Modeling Languages. In: *24th ACM SIGPLAN Conf. Companion on Object Oriented Programming Systems Languages and Applications*, pp. 819–820. ACM (2009). DOI 10.1145/1639950.1640031
74. Vacchi, E., Cazzola, W.: Neverlang: A framework for feature-oriented language development. *Computer Languages, Systems & Structures* **43**, 1–40 (2015). DOI 10.1016/j.cl.2015.02.001
75. Van Der Aalst, W.: Data science in action. In: *Process mining*, pp. 3–23. Springer (2016)
76. Vernadat, F.B.: Enterprise modelling: Research review and outlook. *Comput. Ind.* **122**, 103,265 (2020). DOI 10.1016/j.compind.2020.103265
77. Welsh, M.: The end of programming. *Communications of the ACM* **66**(1), 34–35 (2023). DOI 10.1145/3570220
78. Weske, M., et al.: *Concepts, languages, architectures*. Business Process Management (2007)
79. The ADOxx Metamodeling Plattform. <https://www.adoxx.org> (2023). [Online; accessed 06-August-2023]
80. Introducing Camunda Platform 8. Design, automate, and improve any process across your organization. . <https://camunda.com/platform> (2022). [Online; accessed 06-August-2023]
81. MPS Meta Programming System. Create your own domain-specific language. . <https://www.jetbrains.com/mps> (2023). [Online; accessed 06-August-2023]
82. A free, open-source ontology editor and framework for building intelligent systems. . <https://protege.stanford.edu/> (2023). [Online; accessed 06-August-2023]
83. Zaiane, O.R., Chen, J., Goebel, R.: Dbconnect: mining research community on dblp data. In: *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis*, pp. 74–81 (2007)
84. Zhang, C., Yu, Q., Fan, Q., Duan, Z.: Research collaboration in health management research communities. *BMC Medical Informatics Decis. Mak.* **13** (2013). DOI 10.1186/1472-6947-13-52

Appendix

Number	Target Group	Question	Answer Set
Basics			
2.1	both	Would you consider yourself closer to research or industry?	research or industry, mandatory field
2.2	both	Which is the country, you're currently living in?	selection, mandatory field
2.3	both	What is your current position?	text field
2.4	both	What age group do you belong to?	selection, mandatory field
2.5	both	Which gender do you identify the most with?	selection, mandatory field
Past			
3.1	research	Out of these three, which conference is the most important one for you?	ER, BPM, Models, mandatory field
3.2 (Q1.2)	research	Which are the most important fundamental papers and books for you?	text field (no minimum, max. 10 answers)
3.3 (Q1.3)	research	What are the most important research methodologies you apply?	selection from list (no minimum or maximum)
3.4 (Q1.3)	research	What are the most important research methodologies you apply?	Ordering 17 possible answers (no minimum)
4.1	industry	In what kind of modeling is your company/enterprise most interested in?	software, data, process modeling, mandatory field
4.2	industry	Are there any conceptual modeling conferences in which you are speaking about modeling topics or participating as a listener? If yes, which ones?	No. Yes, for example
Present			
5.1 (Q1.1 & Q1.4 & Q2.2)	both	What are currently the most exciting topics and application areas in modeling for you?	text fields (no minimum, maximum 25 answers)
5.2 (Q1.1 & Q1.4 & Q2.2)	research	As a researcher, for which modeling topics do you see an explicit need for action?	text fields (no minimum, maximum 25 answers)
5.3 (Q1.1 & Q1.4 & Q2.2)	industry	As a practitioner, for which modeling topics do you see an explicit need for action?	text fields (no minimum, maximum 25 answers)
Future			
6.1 (Q1.5)	both	What does modeling have to achieve to be even more relevant in 10 years?	text field
6.2 (Q1.5)	both	Describe your vision for modeling	text field
Community building			
7.1 (Q2.3)	both	The 3 modeling communities (software, data, and process modeling) are closely connected	5-scale
7.2 (Q2.3)	both	The 3 modeling communities (software, data, and process modeling) should be more closely connected	5-scale
7.3 (Q2.3)	both	So why do you think the communities should be more connected? Or why do you think the communities should not be more connected?	text field
Industry-research cooperations			
8.1 (Q2.4)	both	Do you think that modeling research should cooperate more with the industry?	yes/ no because and text field
8.2 (Q2.4)	research	As a modeling researcher, what do you expect from collaborating with industry?	text fields (no minimum, maximum 25 answers)
8.3 (Q2.4)	industry	As a practitioner, what do you expect from modeling research?	text fields (no minimum, maximum 25 answers)
Fin			
9.1	both	Would you like to comment this questionnaire, or would you like to add information for us to better understand your answers?	text field

Table 5 Questions of the survey