CrossMark

Method engineering in information systems analysis and design: a balanced scorecard approach for method improvement

Kurt Sandkuhl[1,2] · Ulf Seigerroth[2]

Abstract

Modeling methods have been proven to provide beneficial instrumental support for different modeling tasks during information system analysis and design. However, methods are a complex phenomenon that include constructs such as procedural guidelines, concepts to focus on, visual representations and cooperation principles. In general, method development is an expensive task that usually involves many stakeholders and results in various method iterations. Since methods and method development are complex in nature, there is a need for a well-structured and resource-efficient approach for method improvement. This paper aims to contribute to the field of method improvement by proposing a balanced scorecard-based approach and by reporting on experiences from developing and using it in the context of a method for information demand analysis. The main contributions of the paper are as follows: (1) It provides a description of the process for developing a scorecard for method improvement; (2) it shows how the scorecard as such can be used as a tool for improving a specific method; and (3) it discusses experiences from applying the scorecard in industrial settings.

Keywords Method improvement · Balanced scorecard · Method engineering · Information demand analysis method

1 Introduction

Modeling methods provide structured guidance for performing complex modeling tasks including procedures to be performed, concepts to focus on, visual representations for capturing modeling results, competences recommended for the modelers, tools and cooperation principles (see Sect. 2.1). Method engineering (ME) is an expensive and knowledge-intensive process, which usually involves many stakeholders, takes a long period of time from first draft to mature method and results in various engineering iterations. This paper mainly aims at contributing to the field of ME within information system analysis and design. We propose a balanced

✉ Ulf Seigerroth
ulf.seigerroth@ju.se

Kurt Sandkuhl
kurt.sandkuhl@uni-rostock.de; kurt.sandkuhl@ju.se

1 University of Rostock, Albert-Einstein-Str. 22, 18059 Rostock, Germany

2 School of Engineering, Jönköping University, Box 1026, 55111 Jönköping, Sweden

scorecard (BSC)-based approach and report on experiences from using the approach and the BSC in the context of improving a method for information demand analysis (IDA).

There is quite a substantial body of knowledge in the field of ME (see Sect. 2.2), but recent research showed that the work on the value of modeling and methods is sparse [1]. Indicator systems and scorecard-based management instruments have been proposed for many organizational functions [2], but in ME there is not much work on quantifying method improvement [3].

The primary perspective for method improvement taken in this paper is that of an organization using the method for business purposes and aiming at improving the contribution of the method to the organization's business objectives. In this context, approaches from the field of Business Value of Information Technology (BVIT) are relevant and were investigated. Most of the BVIT approaches currently existing originated from the demand of enterprises to evaluate the contribution of information technology (IT) artifacts to business success. Section 2.3 includes an overview of BVIT approaches. As modeling methods are IT artifacts and our focus is on organizations using the method as a means to conduct business (e.g., in consultancy projects for other enterprises), BVIT approaches are considered to be relevant.

One of the general approaches for measuring BVIT is to capture indicators for different perspectives of business value in a scorecard with different, balanced perspectives. A scorecard is a data collection tool that helps organizations to reach goals by evaluating progress toward objectives; that is, scorecards commonly help to visualize progress against a goal. Due to the seminal work of Kaplan and Norton [4] in this field, the term "balanced scorecard" is tightly linked to a management instrument that includes several dimensions in an equal ("balanced") way, which is a top-down reflection of a company's mission and strategy. A BSC is future oriented by targeting progress with respect to goals. Furthermore, it is focused on measures that are most critical to the success of the company's strategy.

Our suggestion is to apply this approach for method improvement, as it allows for the combining of different aspects that are relevant for the method's contribution to business objectives. These aspects include the impact of the results achieved by using the method, the understandability and utility of the method documentation, and the guiding power of the work procedures included in the method. The application of the scorecard is illustrated using the IDA method. For organizations using the IDA method in their own projects, the scorecard was supposed to be a management instrument for the operational use of the method.

The main contributions of the paper are (1) a description of the process for developing a scorecard for method improvement; (2) the scorecard as such (as a tool) for improving the IDA method; and (3) experiences from applying the scorecard in industrial settings. The remainder of this paper is structured as follows: Sect. 2 summarizes the foundation for our work from ME and BVIT research. Section 3 introduces the research approach taken, and in Sect. 4, the IDA method is briefly presented. Section 5 describes the development process of the BSC and the resulting "method scorecard." Section 6 is dedicated to experiences related to the use of the scorecard. Section 7 summarizes the results and gives an outlook on future work.

2 Theoretical foundation

2.1 Conceptual base for method engineering

Method engineering needs to be based on a solid foundation. The ME challenge requires a common conceptual base capturing, for methods in general, both methods as a specific instrumental support for actions and the conceptual constituents of methods. In this section, these two conceptual foundations are presented.

All engineering activities are performed to make a difference. During ME activities, purposeful actions are performed by actors aiming at certain ends. According to a model
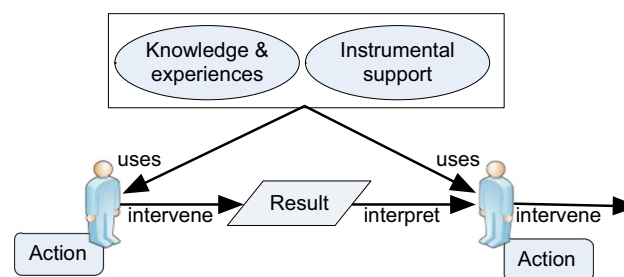

Fig. 1 Support for engineering actions

for socio-instrumental actions by Goldkuhl [5], actions performed by an actor are based on a pre-assessment of grounds, earlier actions performed by the actor(s) and actions producing results directed toward another actor. During the performance of actions, the actor uses instruments, knowledge and experiences as support to make a difference [6, 7] (cf. Figure 1). In this paper, the focused activities are ME and BSC engineering.

Instrumental support for actions can be manifested as different types of artifacts such as methods, theories, tools, patterns and best practices, where previous successful actions have been packaged into prescriptive guidance (instrumental support) for different situations [6].

Methods are widely used as instrumental support for different engineering and development activities in the context of enterprises, such as for Enterprise Modeling (EM), Enterprise Architecture Design and Information Systems Design. According to our view, the use of methods is to be regarded as artifact-mediated actions where different prescribed method actions will guide the development work. We rely on the assumption that a method as an artifact is something that is created by humans and that the artifact cannot exist without human involvement, either by design or by interpretation (cf. [8]). An artifact can therefore be instantiated as something with physical and/or social properties, which also need to be taken into consideration during ME. A method is prescriptive in character, since it gives guidance on what to do in different situations in order to produce certain results and to reach certain goals (cf. [6]). During ME, we can consequently seek support from a method designated for ME (a meta-method). Our focus in this paper is on methods, method engineering and method improvement. We therefore also acknowledge the ISO/IEC 24744:2014 standard and its definition of methods:

specification of the process to follow together with the work products to be used and generated, plus the consideration of the people and tools involved, during an IBD development effort. (ISO/IEC 24744:2014, p. 2)

Based on this definition and Goldkuhl et al. [9], our stance is that a method includes the following constituents:

- *notation* for documentation and representation (documentation rules),
- *procedural guidelines*, tightly coupled to notation, which include meta-concepts such as process, activity, information and object,
- *concepts*, which are the bridging glue between procedural guidelines and notation.

When these three components are closely interrelated, it is referred to as a method component. A method component provides prescriptive knowledge and guidelines for *what* to do and *how* to do a certain task. A method component is also similar to the concepts of method chunk [10, 11] and method fragment [12]. A method component is supposed to give focus to an engineering activity by addressing a focal area. This can be compared with the UML (Unified Modeling Language), which is manifested through two main focal areas, structure and behavior. Each of these focal areas is then divided into a number of sub-focal areas (diagrams) that direct attention toward selected constructs, concepts and their relations, making sense or giving rationality to the specific perspective. Examples are class and object diagrams in UML, which put the focus on concepts or constructs such as class, object, attributes, operations and relations. A method component such as "method engineering" would therefore be executed through the procedural instructions—notation rules—and the related concepts to focus on.

A method is often a compound of several method components into what is frequently referred to as a methodology [13]. A compound of method components together forms a structure that we refer to as a framework. A framework constitutes a phase structure telling us *what* to do, in what *order* to do things and what *results* to produce.

All methods are based on some foundations or perspectives that inform the essence of the method. This method foundation includes values, principles and categories that are manifested in the method and its method components. In UML, for example, the foundation or perspective is object orientation. Object orientation is upheld and constituted through the understanding of encapsulation, polymorphism and inheritance. Consequently, this means that UML must provide a method component(s) that can uphold these principles and where class and object diagrams are perfect examples of this. Encapsulation through these private attributes can only be accessed through the operations provided. The perspective in a method depicts the epistemological, ontological, theoretical and practical standpoints that should be manifested through the method.

The above-mentioned ISO/IEC 24744:2014 definition also depicts people involved in the use of methods for information-based domain development (IBD) efforts. In our perspective, this implies how different people interact and cooperate when performing method-guided work, that is,

collection and cooperation principles. A method component with its procedural guidelines can be used with several different cooperation and collection principles, such as seminars, group work, brainstorming sessions, interviews and questionnaires.

2.2 Method engineering

Method engineering is an expensive and knowledge-intensive process, usually involving many stakeholders and resulting in various engineering iterations. ME is defined by Henderson-Sellers et al. [14] as the engineering discipline to design, construct and adapt methods, techniques and tools for systems development, which also is in line with the IEEE definition of software engineering (SE) and the ISO/IEC 24744:2014 standard. An ME approach that has received much attention is situational method engineering (SME) [14]. In this paper, we have applied a phase-based ME process similar to the ISO/IEC 24744:2014 standard. In this, we have especially acknowledged the iterative interplay between method generation and method validation (enactment) according to Fig. 2. In the ISO/IEC standard, the main activities are generation and enactment. Generation is the act of defining and describing the method based on a defined foundation, often a meta-model. Enactment is the act of validating the method through application. This standard also depicts roles such as the method engineer, who is the person who designs, builds, extends and maintains the method, and the developer, who is the person who applies the method during enactment. These two activities and roles are interlinked so that they can participate during both generation and enactment in an interactive way.

In our ME process for the IDA method, we developed and used a BSC as a tool for method improvement. This was done through generating and measuring different performance indicators in a BSC as part of the validation (enactment) of the method. In doing so, the BSC has also gone through generation and validation as part of the total ME process (see Fig. 2); that is, the generation and enactment of the BSC have been both interwoven in the ME process and parallel activities. This iterative interplay between generation and enactment for both the method and the BSC has called for a structured way to deal with this from the dimensions of both theoretical and empirical input and feedback. For this, we followed the approach of Goldkuhl [15], proposing three levels
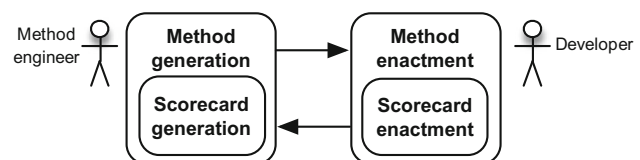

Fig. 2 Method engineering process, based on ISO/IEC 24744:2014

Table 1 Generation and enactment of the information demand analysis method and the scorecard

	Generation	Enactment (validation)
Internally	Reflective discussions between method engineers and developers where the emerging internal structure and content of the method and the scorecard were questioned and developed	Evaluation of method and scorecard consistency in use in terms of structure and interrelationships of its various parts
Theoretically	Use of a method notion to provide a conceptual structure for the method. Relating concept definitions to existing methods and established knowledge (e.g., BSC, BVIT and IDA)	Comparison of the generated method against existing method notions and method theories. Analysis of method and BSC in comparison with existing method and BVIT practices
Empirically	Interview-based investigation for deriving method focus, conceptual foundation and requirements. Practical application of the evolving method and BSC together with industrial partners. Development of support documentation (method and BSC handbook) together with industrial partners	Several test cases for evaluating the usefulness of the method and BSC in industrial cases. Industrial method use and evaluation by external parties by means of a method evaluation framework, NIMSAD. Industrial use and evaluation of method and BSC handbook

to address during method generation and enactment (internal, theoretical and empirical), see Table 1. This approach is similar to other approaches, which also advocate the theoretical and empirical dimensions of ME (cf. [16, 17]). During our ME process, we have performed different generation and enactment activities, illustrated in Table 1.

To our knowledge, research about method validation through the use of BSC to measure ME success for method validation is scarce (see also Sect. 2.4). Some work can be found in relation to the actual design or construction of methods, for example, in Henderson-Sellers et al. [14]. Harmsen [18] presents a more elaborated and promising approach for

using performance indicators (PI) for measuring success in information system (IS) engineering. These indicators are divided into three groups: process-related PI, product-related PI and result-related PI. Even though this research has a focus on IS engineering, we believe that the same principles can be useful for method validation during ME.

2.3 Business value of IT and balanced scorecard

During the last decades, numerous research activities from business administration, economics and computer science have focused on how to measure the business value of IT. Four typical examples of this are as follows:

- Process-oriented approaches, such as IT Business Value Metrics [19]. In process-oriented approaches, the BVIT is demonstrated through process improvements. These approaches investigate how value is added to the business.
- Perceived value approaches, such as the IS Success Model [20]. These approaches base BVIT evaluations on user perceptions rather than on financial indicators or measurements within technical systems.
- Project-focused approaches, including Information Economics [21]. This kind of approach basically tries to support decision making on whether an IT project should be started by calculating a score for project alternatives.
- Scorecard-based approaches, as per the BSC [4]. These approaches try to include different perspectives when evaluating business value, including, e.g., financial, process-oriented and learning perspectives.

As stated in Sect. 1, the focus of our work is on improving the contribution of a method to the business objectives of an organization. All four types of BVIT approaches could potentially be tailored for this purpose. However, there are differences between these approaches with respect to their suitability:

- The required method improvement approach has to include business value and coherence with business drivers such as reduced lifecycle time or increased flexibility. These business drivers are measurable criteria reflected in the control systems of many companies. Perceived value approaches do not cover these aspects sufficiently.
- Method improvement requires monitoring of relevant indicators during a longer period of method use, i.e., capturing of performance indicators only once would not be sufficient. This requirement is hard to meet with project-centric approaches.
- Process-oriented approaches are by nature quite specific to the individual company, as they require an understanding of business processes, the potential business impact and the potential IT impact before starting the actual analysis

of BVIT. This makes these approaches quite expensive for method improvement in terms of efforts to be invested, as methods are expected to be used in many different organizations.

Among the scorecard-based approaches, the BSC proposed by Kaplan and Norton [4] is the most established. The BSC is a *management system*; that is, it includes measurement approaches to continuously improve performance and results.

2.4 Scorecard use in quality management

Scorecard-based approaches for BVIT evaluation have their origin in research in economics and business administration and thus are outside the traditional scope of SE and IS development. However, there are links between the field of quality management [22, 23] and BVIT evaluation; these will be analyzed in this section. The links include:

- earlier applications of scorecards in Quality Management (QM),
- the use of QM approaches within scorecards and
- the general organization as management systems.

2.4.1 Management systems

Management systems commonly enable an organization to systematically develop certain features in a continuous process [24]. When implementing management systems, organizational roles are established, procedures defined, documentation standards specified and organizational policies and cultures or mindsets set up. The actual management task in a management system usually includes continuous improvement cycles [25]. Quality management systems in organizations and the scorecard-based evaluation of BVIT are both by nature management systems, in that they require a certain organizational setup as described above. They also do not only comprise solitary assessments or data collection, but continuous tasks integrating many assessment activities into an overall picture of the situation and its development. As scorecard-based approaches are considered "lightweight" management systems [4], the difference to QM systems is substantial if actual implementations are compared. Scorecards are supposed to focus on a handful of goals and three to five perspectives [4]. QM systems commonly cover the complete lifecycle of a software product, information system or IT system, with procedures and instructions in place for most kinds of artifacts developed during this process.

2.4.2 Applications of scorecard in QM

Scorecard-based approaches have been used in only a few application scenarios in SE and IS development. This was the result of a literature survey we performed in February 2018. For the literature survey, we followed the guidelines of Kitchenham [26]. When planning the literature review, we started by defining the research question to be tackled as follows:

In the field of software and information systems development, which scientific publications exist on the development and application of scorecards for quality management?

Furthermore, we decided to include IEEE, ACM, Springer Link and Scopus as publication outlets. The assumption we made is that all important developments relevant to the research question should be visible in these outlets. Relevance to the research question has to be based on specifying inclusion and exclusion criteria. In our case, we developed a search string to be applied on the publication outlets:

"scorecard" AND "quality" AND ("management" OR "assessment" OR "control" OR "assurance") AND ("software" OR "information system" OR "IS" OR "IT-system" OR "IT artefact" OR "IT-artifact")

In the search string, we took into account that synonyms or related terms for "quality management" (i.e., quality assessment, assurance or control) or for "software" and "information system" might be used (i.e., IT-system, IT artefact/artifact or IS). The result of the search with the above search string in title, abstract and keywords was a total of 762 papers with a number of duplicates, as Scopus also includes part of IEEE and ACM. Table 2 shows the number of hits for different parts of the search string. Among the papers was also our work that formed the starting point for this paper [27]. This work was excluded from the analysis.

In the next step, we checked the relevance of the papers by reading the abstracts. Many of the papers had to be excluded because they were from application domains using scorecards and information systems (e.g., hospitals, metallurgy, supply chains, e-government) for quality-related activities, but were not concerned with the actual quality of software or information systems. Other papers were excluded because they focused on generating or improving the quality of scorecards by using, for instance, ontologies, self-organization or machine learning. This first step reduced the number of papers to 33. In a second step, the full text of the remaining papers was checked for relevance. Again, a number of papers had to be excluded because they were focused on project management issues, price evaluation of software or general management systems. After the second step, only five publications were left.

In total, the survey revealed the following publications on scorecard use:

Table 2 Overview of the results of the literature search

Search string	No. of hits	Analyzed in step 2
scorecard AND (quality AND assessment) AND software	13	1
scorecard AND (quality AND assurance) AND software	12	1
scorecard AND (quality AND management) AND software	53	11
scorecard AND (quality AND control) AND software	28	2
scorecard AND (quality AND control) AND (information AND system)	88	1
scorecard AND (quality AND assurance) AND (information AND system)	33	3
scorecard AND (quality AND assessment) AND (information AND system)	63	2
(scorecard) AND (quality AND management) AND (information AND system)	191	13
["scorecard" AND "quality" AND ("management" OR "assessment" OR "control" OR "assurance") AND ("software" OR "information system" OR "IS" OR "IT-system" OR "IT artefact")]	762	33

- Sivaji and Tzuaan investigate scorecard use for improving Web site usability and User eXperience (UX) [28].
- Staron et al. [29] propose a scorecard for managing code stability indicators in the context of monitoring the performance of software development organizations.
- Chang and King [30] develop a functional scorecard for information systems addressing business process effectiveness and organizational performance. This scorecard includes a number of proposals on which factors and indicators should be considered.

- Keyes proposes an IT scorecard including perspectives, criteria and indicators for the IT in an organization [31].
- Subramanian et al. [32] investigate critical factors in IS implementation strategy, software quality and software project performance.

None of the above publications investigates the use of scorecards in ME. An additional literature search on scorecard, method and engineering, improvement, evaluation and validation returned only our paper [27] as a result.

2.4.3 Use of quality management approaches within scorecards

The clearest and probably most important link between QM and scorecards is the use of measurement approaches or metrics from QM for determining the indicator values used in scorecards [33]. Primarily, if the business goals of an enterprise include quality goals and, as a consequence, the business value depends on the quality, the indicators used for capturing goal achievement and business value will be quality related or even originate from QM. However, the perspective of the scorecard still remains different from the QM view. In our approach for BVIT evaluation presented in Sect. 4, some indicators originate from QM, some are inspired by QM experiences and others are from economics. When discussing the different perspectives of the scorecard and their indicators, this QM relation is elaborated in more detail (see Sect. 5.4).

Usability questionnaires and test checklists are not considered as scorecards since their focus is on assessing characteristics and features of artifacts and systems rather than evaluating business value. However, such instruments can be applied for determining indicators contributing to the business value investigation (see also the next section).

3 Research approach

The research approach for the development of the BSC and the engineering of the IDA method combines Design Science (DS) [34] and Action Research (AR) [35]. In this combination, we have taken a stance in Technical Action Research (TAR) according to Wieringa and Morali [36]. In TAR, the engineering process and the artifact design are the starting point and are where the artifact is supposed to be validated in practice in a scaled-up sequence. It starts with testing a prototype (artifact) in a controlled environment, via testing of the prototype in context (real-life setting) to full-fledged applications of the final artifact to solve a group of real practical problems in an enterprise [36]. Our artifacts in this case are twofold: The first one is the IDA method, which is developed as a "treatment" to improve or solve information challenges
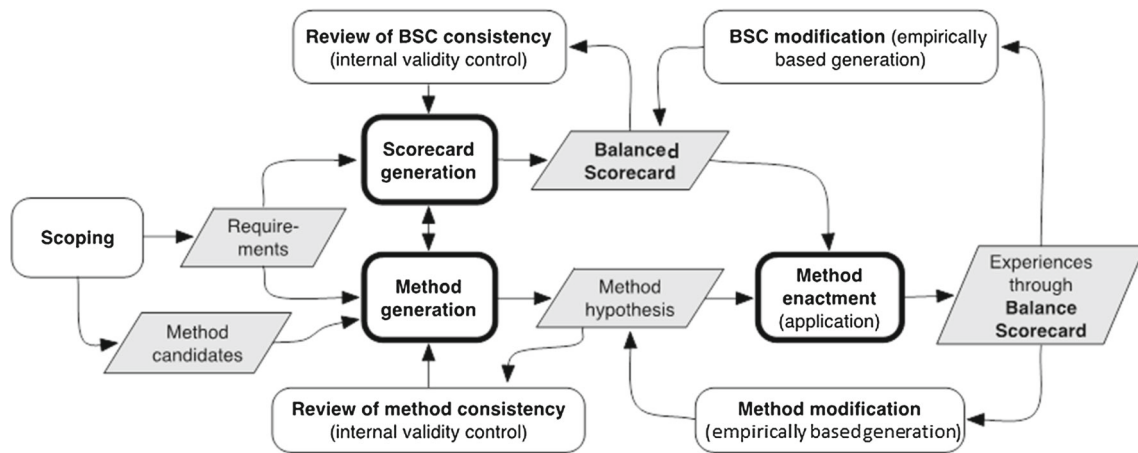
Fig. 3 Generation and validation of the information demand analysis method via the balanced scorecard

in enterprises in dimensions of information supply, provision, demand, logistics and so on. The second artifact is the BSC, which is developed as a "treatment" for method improvement. Through this approach, the artifacts are motivated by the desire to solve a *class* of problems, rather than specific problem instances, which is an important distinction and a guiding principle of TAR. There are also earlier promising initiatives to combine DS and AR, one example of which is Action Design Research (ADR) by Sein et al. [37]. Even though the artifact is in focus in ADR, the approach is still problem driven [38]. In our case, the TAR approach is more convenient since the method is developed from the notion that we have to handle information challenges in enterprises, while the BSC relies on the notion that there is a need for method improvement.

The TAR approach also emphasizes the co-production between research practices; that is, the development of the artifacts was performed through continuous improvements in collaboration with our case partners. This was mainly done through the researchers taking on three different roles: *Designer*, *Helper* and *Researcher*. The role as designer covers the actual design of the initial technique (artifact), which, as a final goal, should be able to support the solving of a class of problems. The helper employs the usage of the technique (artifact) to help others (case partners). And, of course, there is the role of the researcher, drawing lessons learned about the capabilities of the artifact. The artifacts have emerged from the interaction with the organizational context, even if the initial design was guided by the researchers' intent. Consequently, the artifacts were shaped by the organizational context during their development and use. In fact, both ME and the development of BSC have required a combination of competences from both academia and industry in order to build a joint understanding of the problems and ongoing processes. On a high level, the ambition was that both the IDA method and the BSC would be continuously deployed as pro-

totypes, making it possible for domain experts working at the case partners to evaluate the performance in their production environment. This evaluation and analysis have created valuable feedback for the researchers, thus establishing a virtuous cycle consisting of problem generation, solution, deployment and evaluation—very similar to the overall organization of the TAR methodology.

The purpose of the IDA method was to support identification, modeling and analyzing information demand as a base for the development of technical and organizational solutions providing a demand-driven information provision. The ME process for the IDA method and the BSC is described in Fig. 3.

In this study, we focus on the enactment phase and the use of BSC as a tool for method improvement. Five different IDA cases were included:

- A metal finishing company (coordination of quality, technology and production).
- A municipality (handling of errands).
- The association for Swedish SAP users (coordination of information flow).
- A timber company (identification of information demand for test strategy).
- A gardening retailer (information demand for different organizational roles).

4 The IDA method and its constituents

The IDA method was mainly developed from 2006 to 2008 in the research and development (R&D) project InfoFLOW, which included seven industrial and academic partners. Thereafter, the method has been applied and further developed through several industrial applications both in Sweden and in Germany. The basis for the method development is
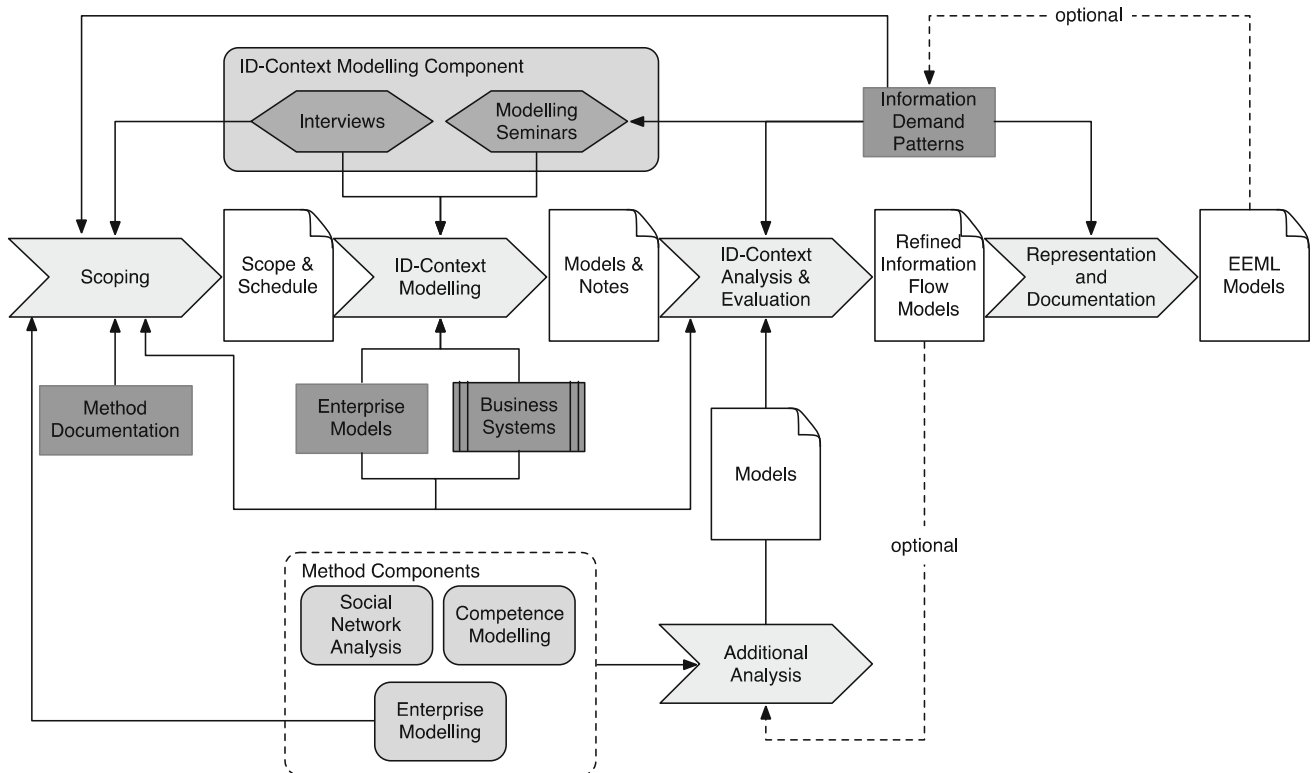
Fig. 4 An overview of the information demand analysis method [42]

the understanding of the term information demand, three industrial cases from the InfoFLOW project [38–40], the requirements derived from these cases [41] and the notion of methods as presented in Sect. 2.1. Since information demand has a close relation to roles (as actor) in a practice, we need to ensure that the designated roles also are represented during the information demand modeling sessions; that is, interactive research guided by the cooperation principles according to the notion of method. Therefore, it is important that the actual modeling is performed together with the actual roles that are to be modeled.

Understanding information demand requires understanding of information demand contexts in terms of a number of dimensions. In Fig. 4, an overview of the IDA method (framework) is presented. Since context is considered central to IDA, method support (method component) for modeling such a context is at the core of the IDA method (cf. [42]). However, in order to be able to perform any meaningful context, modeling a clear scope is needed. Consequently, the IDA starts with scoping activities. Also, depending on the requirements and needs relevant for the specific case, additional aspects of information demand might be analyzed and modeled.

Scoping as a prerequisite for information demand context modeling is the process of defining the area of analysis and the selection of the parts of the organization to be the subject

of the analysis. This phase also includes the identification of the roles relevant for the IDA. Scoping also sets the scene for identification and understanding of the organization's problems, goals, intentions and expectations to motivate it to engage in the IDA.

Information demand context modeling is preferably performed through participative activities such as collaborative modeling seminars, where the participants with their domain knowledge themselves are involved in the actual manufacturing of different models. This process is also supported and facilitated by a method expert, who could be an internal or external person. As illustrated in Fig. 4, the conceptual focus in this phase is on information demand within a defined scope. The key to context modeling is to identify the interrelationship between roles, tasks, resources and information. No direct attention is paid to the sequence of activities, resource availability and so on.

Once the necessary knowledge about the information demand contexts is defined, it is used for a number of different purposes. One purpose is evaluation, where different aspects of information demand are assessed in relation to roles, tasks, resources and information. It is also used to address the results from the modeling session with respect to the motivation and purposes expressed during the scoping activities. Focusing on information demand contexts only provides an initial view of information demand. No consideration is given

to aspects such as individual competence, organizational expectations and requirements in terms of goals, processes and so on. Depending on the intentions behind the analysis, further activities will be required. The IDA method provides a number of method components supporting such activities. Since the main focus of the IDA method presented here is on information demand, it utilizes existing procedures and notations for such additional aspects rather than defining new ones.

In the IDA method, patterns also have a significant role. For the conceptual positioning of patterns, see Sect. 2.1. The general idea of information demand patterns is similar to most pattern developments in computer science: to capture knowledge about proven solutions in order to facilitate the reuse of this knowledge. In this context, Martin Fowler's statement "A pattern is a solution to a problem in a context" serves as inspiration. For our work in the InfoFLOW project, we have defined the term information demand pattern as follows:

> An information demand pattern addresses a recurring information demand problem that arises for specific roles and work situations in an enterprise, and presents a solution to it. [43]

An information demand pattern is constituted by a number of essential parts used for describing the pattern: (1) A statement about the *organizational context* where the pattern is useful. This statement identifies the application domain or the specific departments or functions in an organization forming the context for the pattern definition. (2) The *problems* of a role that the pattern addresses. The tasks and responsibilities a certain role has are described in order to identify and discuss the challenges and problems that this role usually faces in the defined organizational context. (3) The *solution* that resolves the problem, which for information demand patterns includes three parts:

- The *information demand* of the role, which is related to the tasks and responsibilities and is described as part of the pattern, i.e., the different parts of the information demand are identified.
- The *quality criteria* for the different parts of the information demand include the general importance of the information demand part, the importance of receiving the part completely and with high accuracy, and the importance of timely or real-time information supply.
- A *timeline* indicating the points in time when the different information parts should be available at the latest.

The *effects* that participate in forming a solution are manyfold. If the needed information part is not available or arrives too late, this might affect the possibility of the role completing its task and responsibilities. The effects described

in the pattern include potential economic consequences, time/efficiency effects—that is, whether the role will need more time for completing the task or will be less efficient—effects on improving or reducing the quality of the work results, effects with respect to the motivation of the responsible role, learning and experience effects and effects from a customer perspective.

Additionally, a pattern can also be represented as a *visual model*, for instance, as a kind of enterprise model. This model representation is then supposed to support the communication with potential users of the pattern, as it illustrates the information demand context. This includes the relation of the role to co-workers and other roles in the organization, the relation between the different parts of the information demand and IT system in the enterprise, which are potential sources of this information, and the relation of tasks and responsibilities to processes in the organization.

5 Development of BSC for method improvement

The BSC development is illustrated using the IDA method (see Sect. 4). Among the users of the IDA method are consultancy companies who perform many projects aiming at improved information flows in small- and medium-sized enterprises. They consider the method as a kind of resource in their "production" process. These companies are interested in having control of the use of the method from an economic perspective, to find improvement potentials and to get at least an idea of the value for their business.

This means that the primary perspective for method improvement taken in our work is that of an organization using the method for business purposes and aiming at improving the contribution to business objectives. Such an organization could be, for example, a consultancy company offering analysis and optimization services for its clients based on the method, or enterprises using the method internally for detecting and implementing change needs. In an organizational context, improvement processes are usually guided by defined goals and instruments to supervise the goal achievement. The approach proposed in our work is to apply the principles of a BSC for creating a management instrument for method improvement. That is, we will not use the original BSC perspectives and content proposed by Kaplan and Norton [4]. Instead, we will use the process of developing such a BSC for method improvement and the general BSC structure of goals, sub-goals, indicators and so on. Section 5.1 describes the process of developing the scorecard, whereas Sect. 5.2 introduces the actual "method scorecard." Section 5.3 reports on the evolution of the method scorecard.
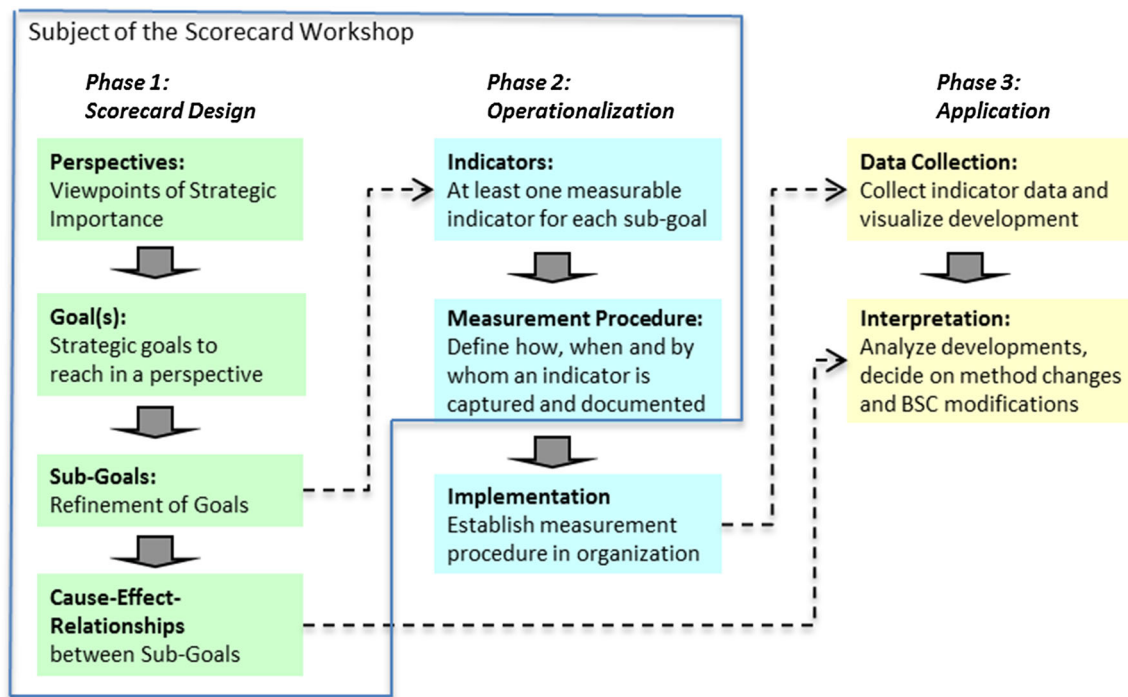
Fig. 5 Phases and steps of method scorecard development. Arrows between the steps of the phases indicate their sequence. Dashed lines with arrows indicate input to other phases

5.1 Scorecard development process

The main instrument for developing the scorecard was a workshop with all organizations planning to use the IDA method. The workshop produced an initial version of the scorecard, which formed the basis for refinements and further development during the project. During the scorecard workshop, the two phases of scorecard design and operationalization were conducted, both of them consisting of several steps. A third phase of scorecard application followed (see Fig. 5).

5.1.1 Scorecard design

In the phase of scorecard design, the following steps were taken. The first step was to evaluate whether the *perspectives* proposed by the original BSC approach (i.e., financial, internal business process, learning and growth, customer) are valid and appropriate for method improvement or should be changed. Starting points for identifying relevant perspectives were the strategic aims of the participating organizations. Before the workshop, the participants were asked to define the strategic aims for the organization they were representing. In the workshop, each organization presented its aims, which were then sorted into groups, and every group was labeled with a suitable headline as the basis for perspectives in the scorecard. The result of this step was an initial agreement on perspectives to consider in the "method scorecard."

For each perspective, *strategic goals* had to be defined and preferably quantified, as quantifying them helps to reduce the vagueness in strategic goals. Identifying strategic goals was again based on the organizations' strategy and aims. The workshop participants agreed to focus on aims that were directly related to the IDA method use. This made it possible to agree on joint aims among all participants, which turned out to reflect the majority of all aims of the individual organizations. The defined strategic goals were in a next step broken down into *sub-goals*. The objective was to define not more than five to seven sub-goals per goal. The last step related to scorecard design was the identification of *cause–effect relationships*. There might be strategic goals or sub-goals that cannot be achieved at the same time because they have conflicting elements. It is important to understand these conflicts or cause–effect relations. Sub-goals and cause–effect relationships were jointly developed by all workshop participants.

5.1.2 Scorecard operationalization

After having covered the scorecard design, the focus of the workshop shifted to operationalization, which is Phase 2 of the scorecard development. For each sub-goal defined in the different perspectives, a way had to be found to measure the current situation. For this purpose, *indicators* had to be defined contributing to capturing the status with respect to

the sub-goal. When defining indicators, one had to keep in mind that there must be a practical way to capture these indicators. In this context, existing controlling systems or indicators (e.g., from quality management) were inspected and checked for possibilities to reuse information. For each indicator identified, the *measurement* or recording procedure was defined. A measurement procedure typically includes the way of measuring an indicator, the point in time and interval for measuring, the responsible role or person performing the measurement and how to document the measured results.

In the workshop, the process of identifying indicators and measurement procedures was to start by discussing, for each sub-goal, which criterion or criteria would be relevant for the evaluation progress of goal achievement. Based on these criteria, potential indicators related to these criteria were identified. This process did not lead to fully specified criteria and an indicator set for all sub-goals, but only to an initial and incomplete list. As follow-up to the workshop, the academic partners were asked to make proposals for additional indicators based on earlier experiences or existing scientific work (cf. Section 5.4). These proposals were discussed in a second working session and resulted in a list of indicators and measurement procedures.

Based on the indicator definition and the measurement procedures, the implementation in the organization had to follow. This step was not part of the workshop, but the workshop resulted in recommendations for the different organizations on how the implementation should be performed. Furthermore, all organizations agreed to use the same forms for collecting the data (cf. Section 6).

5.1.3 Scorecard application

The third phase of scorecard development, application, concerns the use and the further development of the method scorecard during the method improvement process. During this phase, the implemented measurement procedures and the aids developed for collecting data were used in each participating organization. Experiences from data collection and the interpretation of collected indicator values are the subject of Sect. 5.

5.1.4 Role distribution

After the method scorecard development workshop and thus for the operational phase, a role distribution was agreed upon between the participating organizations. This role distribution consisted of the following roles, which basically emerged more from practical considerations than from formal or legal conditions:

- Method manager: the organization taking care of collecting all method improvement proposals, maintaining the

documentation and coordination of the method improvement process. In other contexts, this role might also be called the method owner, but for the IDA method the legal owners were all four organizations participating in the development.
- Method user organization: academic and industrial companies using the IDA method for their modeling projects.
- Method users: individual modelers using the whole IDA method or parts of it.
- Method developer: individual researchers or engineers who were part of the method development process. The method developers all came from one of the four academic or industrial InfoFLOW project partners.
- Scorecard manager: the organization taking care of collecting all scorecard improvement proposals, maintaining the scorecard documentation and coordination of the scorecard improvement.
- Scorecard owner: the legal owner of the method scorecard. In this case, the scorecard is published.

5.2 Scorecard for method validation

The development process described in Sect. 5.1 resulted in four different perspectives in the scorecard, which reflected the shared organizational goals of the enterprises participating in the InfoFLOW project and the development of the method scorecard. Figure 6 illustrates the four perspectives with their strategic goals, which were as follows:

1. Method Documentation Quality: The quality of the IDA method handbook and aids

   Goal: To have a method that is easy to train and communicate

2. Pattern Quality: The quality of the Information Demand (ID) patterns

   Goal: To achieve patterns of high quality applicable with the method

3. Process Efficiency: The efficiency of the process for understanding information flow problems in enterprises and developing an appropriate solution proposal

   Goal: More efficient processes and resource use for the analysis, including a proposal for a solution

4. Solution Efficiency: The efficiency of the solution implemented in an enterprise based on using the IDA method and ID patterns
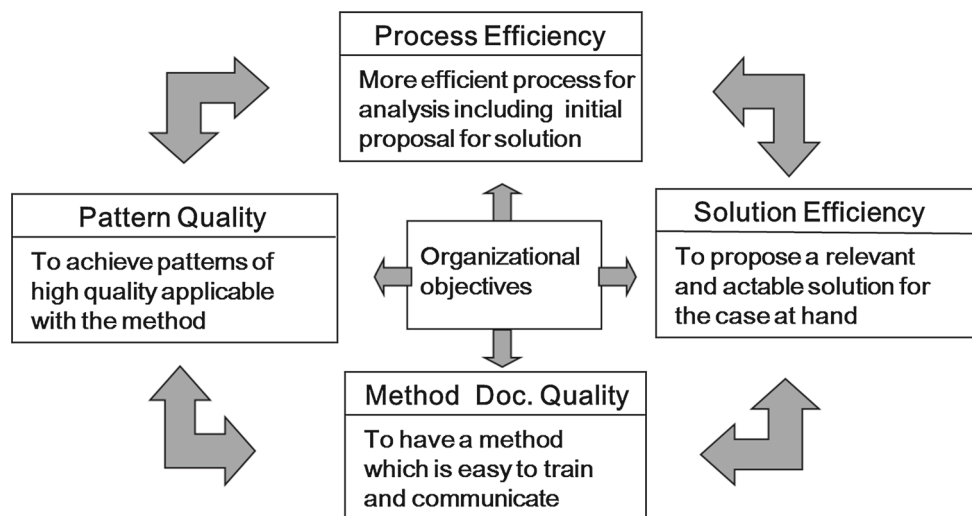
Fig. 6 Perspectives and main goals of the method scorecard

Goal: To propose a relevant and actable solution for the case at hand.

5.2.1 Sub-goals

The sub-goals for each perspective served as refinement of the goals and as a first step to operationalizing the goals. The sub-goals for the different perspectives and their goals are presented in Table 3.

5.2.2 Cause–effect relationships

The identification of cause–effect relationships started at the level of perspectives; that is, if juxtaposing the different perspectives, are there any goals or sub-goals in one perspective contradicting goals or sub-goals of another perspective? The contradicting goals/sub-goals were important to discover, because such a contradiction would basically mean that one sub-goal hinders achieving another one. This kind of relationship was not discovered in the scorecard development process. However, there are a number of goals/sub-goals supporting goals or sub-goals of other perspectives. For example, high pattern quality is expected to support solution efficiency, as patterns are meant to capture proven and reusable parts of solutions. Also, high documentation quality is supposed to support process efficiency, because well-documented process steps should lead to reliable execution of these steps. An awareness of this kind of supportive relationship helps understanding potential side effects when planning method improvement measures.

5.2.3 Indicator examples

Our view of a method being a guide for actions, which often are artifact-mediated actions (see Sect. 2.1), affected the selection and definition of indicators. For brevity reasons, we will only present an excerpt of the criteria and indicators of the method scorecard. These criteria and indicators are captured for each perspective in a separate table, including the following information:

- What to measure, i.e., the criteria to capture. Criteria are grouped into aspects.
- Motivation of these criteria and comments (not included in Tables 4, 5, 6 and 7).
- Indicators reflecting the criteria. This is the actual value to measure.
- Indicator description. Explanation related to the indicator name.
- Practical implementation of capturing the indicators, i.e., how to measure, who will be responsible for measuring, when to measure and how to document the findings.

The complete method scorecard has 61 indicators; all the indicators are presented on 18 pages and are available in a technical report from the InfoFLOW project [1].

Table 4 shows an excerpt of the criteria and indicator table for the perspective "method documentation quality." This excerpt is focused on "documentation quality." Further aspects in this perspective are method documentation maturity, method support for pattern use and method support for pattern extension.

An excerpt of the criteria and indicators for the "pattern quality" perspective is given in Table 5. This excerpt relates

Table 3 Sub-goals for the four perspectives and goals of the method scorecard

Perspective	Goal	Sub-goals
Method documentation quality	To have a method that is easy to train and communicate	Easy to teach the method and train future modelers (transferability)
		Provide good documentation
		Method shall support the effective development of new patterns
		Method shall take into account that patterns need continuous improvement
Pattern quality	To achieve information demand patterns of high quality applicable with the method	Applicability of patterns, e.g., it should be easy to decide whether a pattern is applicable or not and easy to actually apply the pattern
		Support for deciding which parts of a pattern to apply
		High "technical quality" (quality of details, visualization, consistency of description)
		Positive contribution to the method for information demand analysis
		Relevance for different application domains or explicitly domain-independent pattern
		Possibility to improve the pattern or to further develop it
Process efficiency	More efficient process for information demand analysis including initial proposal for a solution	To facilitate faster processes
		To reduce effort (less expensive processes)
		High quality of the solution developed
		Clarity of results and next steps
		Relevant solution, easy to understand
		Feasible solution
		To have a process that is easy to understand
		For the customer
		For the consultant/analyst applying it
		High security that a solution can be delivered (robustness)
Solution efficiency	To propose a relevant and actable solution for the case at hand	Include all phases of developing the solution (e.g., analysis, design, prototype, implementation)
		High quality of the solution from the customer perspective
		Clarity of results and next steps
		Relevant solution, easy to understand
		Feasible solution

Table 4 Excerpt from criteria and indicators for the "method documentation quality" perspective

Aspect	Criterion	Indicator name	Indicator description	How to capture indicators
Documentation quality	Learning time	Average learning time for new analyst until "productivity"	How much time does it take on average until a person can be considered "productive"?	Captured during training sessions by the method specialist
		Average learning time for new trainers until "productivity"	How much time does it take on average until a new trainer for the method can be considered "productive"?	
		Average learning time for participants in analysis projects until able to participate in analysis projects	How much time does it take on average until a participant in analysis projects understands what she/he is contributing to?	
	Perceived quality of method documentation	Perceived quality of completeness, correctness, understandability, etc. on a suitable scale (e.g., five-point Likert scale)		
	Documentation of maturity	Maturity level according to review status of different stakeholder groups, such as consultant, method specialist or modeling facilitator	What maturity level between "draft" and "fully validated" is assigned in method reviews by domain experts?	
		Number of improvements proposed by the researchers	How many change requests were submitted for the method by the researchers?	
		Number of shortcomings detected in use	How many change requests were submitted by practitioners?	

to the aspect of applicability. Other aspects covered are technical quality and extensibility.

Table 6 shows the criteria and indicator selection for the third perspective, "process efficiency," with a focus on the aspect of the analysis process. Other aspects include indicators for the analysis result (solution) and the delivery of the solution.

An excerpt of the criteria for the perspective "solution efficiency" is presented in Table 7. This excerpt focuses on strategic benefits. Other aspects are automation benefits and transformation benefits.

Table 5 Excerpt from criteria and indicators for the "pattern quality" perspective

Aspect	Criterion	Indicator name	Indicator description	How to capture indicators
Applicability	Learning time	Average learning time for new team members until "productivity"	How much time (in hours, not calendar time) does it take on average until a person understands the pattern and can explain its use?	Captured during pattern use by method specialist
		Average learning time for existing team members until "productivity" with a new pattern	How much time (in hours, not calendar time) does it take on average until a person familiar with the pattern concept understands the new pattern and can explain its use?	
	Intensity of use	Average number of uses	How many times has the pattern been used in projects?	
		Perceived applicability of patterns when reusing them	How is the applicability of a pattern rated on a scale between −2 and +2?	
	Domain independence	Number of application domains expected to be relevant for a pattern	According to the pattern developers, how many application domains exist for the pattern?	
		Number of application domains where a pattern was used and average number of applications	In which domains was the pattern actually used? Average number of applications in each domain?	

5.3 Evolution of the method scorecard

The method scorecard as such did not only serve as an instrument for MI, but the scorecard itself was also subject to several improvement steps. The first version presented in Sect. 5.2 was used by two industrial and two academic method user organizations that cooperated in developing and improving the IDA method. These four organizations in some cases also performed the IDA projects as joint ventures and they met every 2 months to discuss experiences with the IDA methods and to share practical recommendations in method scorecard use. More information about the method use and the experiences collected are provided in Sect. 5.1.

The first major revision of the method scorecard happened after 1 year of scorecard use based on the experiences collected. This revision no longer included "solution efficiency" as a perspective, because the collection of the indicators required for this turned out not to be implementable in practice (cf. Section 5.1). The core problem experienced was that for solution efficiency a comparison of earlier costs or time frames for information supply and quality problems—that is, the baseline—with costs and time frames after optimization would be required. The organizations optimizing the infor-

Table 6 Excerpt from criteria and indicators for the "process efficiency" perspective

Area	Aspect	Indicator name	Indicator description	How to capture the indicators who measures?
Analysis process	Duration of processes	Average time for performing an information demand analysis	How long does it take to perform an information demand analysis in a project (calendar time)?	
		Average time for scoping phases	How long does it take to perform the scoping phase in a project (calendar time)?	
		Average time for context modeling	How long does it take to perform context modeling in a project (calendar time)?	Capture start date and end date in order to be able to separate different iterations of phases and overlaps between phases; must be captured in days
		Average time for information demand context analysis	How long does it take to perform information demand context analysis in a project (calendar time)?	
	Cost of processes	Average number of hours spent for the information demand analysis	How many hours of consultants are needed for the information demand analysis?	
		Average number of hours spent on scoping phases	How many working hours does it take to perform the scoping phase in a project?	
		Average number of hours spent on context modeling	How many working hours does it take to perform context modeling in a project?	Capture start date and end date in order to be able to separate different iterations of phases and overlaps between CM and CA; duration must be captured in days
		Average number of hours spent on information demand context analysis	How many working hours does it take to perform information demand context analysis in a project?	
		Average costs additional to the hours spent for the information demand analysis	How much in the project was spent on traveling, hotels etc.?	

Table 7 Excerpt from criteria and indicators for the "solution efficiency" perspective

Aspect	Criterion	Indicator name	Indicator description	How to capture the indicators who measures?
Information supply	Quality of information supply	Average number of events/roles with incomplete information supply	How often does a role detect that relevant information has not been provided?	All roles that were part of the improvement process have to do self-recording of cases
		Average number of events/roles with information provided too late	How often does a role detect that relevant information has been provided later than the task at hand required it?	
		Average number of events/roles with incorrect or outdated information provided	How often does a role detect that information provided contained incorrect or outdated content?	
	Cost of information supply problems	Average time for completing task	How much time does a task take where the information supply was improved (on average per role)?	Recording of time by role performing the task or by logging mechanism; comparison with earlier values
		Number of quality problems caused by information flow problems	How many defects or deficiencies were caused by information flow problems (for an organizational unit)?	Quality problem are recorded in quality management system; extension for capturing cause of the problem required
		Costs of quality problems caused by information flow problems	What were the costs of the quality problems caused by information flow problems?	Quality problem are recorded in quality management system; extension for capturing cost of tackling problem required
		Average number of coordination problems due to information flow problems	How many changes in production plans or work assignments had to be made due to missing information (on average per role)?	Recording of occurrences by role performing the task; comparison with earlier values

mation flow were not willing to provide these figures or they were not able to capture them.

The second revision followed after 2 years of scorecard use and concerned only indicators; that is, there were no changes with respect to perspectives, goals or sub-goals. Examples of indicator changes were that the number of previous method uses was no longer captured as this indicator seemed no longer relevant; the indicators for documentation quality were only captured by new method users or after major changes in the documentation; and new indicators regarding training time were added when the documentation of previous IDA method use cases became available for training purposes.

All these major changes were proposed by method user organizations, collected by the method manager and integrated into the third release.

5.4 Origin of scorecard indicators

As indicated in Sect. 2.4, some of the indicators in the method scorecard were based on QM approaches and experiences and others on approaches from economics. This section briefly discusses the origin and theoretical foundation of the indicators, which is summarized in Table 8.

Most indicators in the Method Documentation Quality perspective are based on the work on document quality indicators, in particular [5, 44, 45]. Although these publications take different perspectives, they all address criteria for the general comprehensibility of a text, besides its legibility. Among the criteria, the possibility of extracting information about certain facts and circumstances from a text is addressed. Langer et al. [44] present 18 evaluation criteria, which can be summarized to four attributes: simplicity, structure and alignment, brevity and conciseness, as well as inspiring additions. These criteria and attributes inspired the definition of criteria and indicator in the scorecard.

The indicators on Pattern Quality are based on work and experiences in the area of pattern quality from ontology engineering, pattern quality from enterprise modeling and conceptual model quality. These three areas are considered relevant because information demand patterns are a domain-specific conceptual model from their structure, but also meet typical characteristics of patterns in computer science (cf. [46]). More concretely, we adapted the indicators identified and tested for task patterns [47] for the applicability of patterns, the positive contribution to the method and the possibility of improving the patterns; and the indicators from ontology pattern quality [47] for support for selecting the right patterns, technical quality of patterns and relevance for application domains. Conceptual model quality [48] is a basis for technical pattern quality and is incorporated into the work on ontology pattern quality.

Criteria and indicators for Process Efficiency are well researched in the field of business process management. What to measure (e.g., duration of complete processes or single tasks, effort required, quantity of output, etc.) and how to measure this (automated logs in IT systems, observations, self-recording, etc.) are basically textbook knowledge (see, e.g., [49, 50]) and were applied in combination with indicators from task pattern use [47] and economic aspects of method evaluation [51].

The indicators for Solution Efficiency originate from the work in economics on the effects of IT on organizational transformation. More concretely, the work by Gregor et al. [52] proposes to investigate the strategic, informational, transactional and transformational perspectives. The most relevant perspectives for our work used for indicators are the informational and transactional ones. Informational benefits can be assumed if improvements in the information infrastructure for control, planning or other management tasks are achieved; that is, there is an informational advantage due to the new methodology as compared to the situation without it. Transactional benefits are typically connected to the automation or at least the semiautomation of tasks within an enterprise. This type of benefit is usually related to cutting costs and reducing time required for processes or tasks.

6 Method scorecard in use

The method scorecard was applied in two different contexts: for improving the IDA method and for the enterprise modeling method 4EM [40]. When applying the scorecard in the context of the IDA method, two groups of method users have to be distinguished:

- Members of the method development team. This group obviously consisted of experts in IDA and focused on finding method improvement potential,
- Method users from outside the development team who received training in IDA and used the method on their own shortly after the training. This group is expected to have a more independent perspective on the utility of the IDA method.

Data collected for these different groups are discussed in Sects. 6.1 and 6.2. In order to investigate whether the method scorecard would also be suitable for other methods than IDA, the scorecard was applied in a few 4EM cases (see Sect. 6.3).

6.1 Scorecard use by IDA method developers

In total, four different members of the method development team used the scorecard during five different IDA cases in a time frame of 10 months. The cases addressed informa-

tion flow problems in a municipality and in enterprises from retail, an automotive supplier, a wood-related industry and the IT industry. In each case, several modeling activities were performed, scorecard data collected and observations noted down. The observations were discussed with the other members of the method development team. Several adjustments were made as a result of the observations when using the scorecard, all of them in the first 6 months of scorecard use:

- Initially, data capture in the cases happened based on a printed version of the document describing the scorecard. Since entering this handwritten data into a spreadsheet was tedious, a software tool was developed for data capture. This tool offered the possibility of capturing experiences and remarks in free-text form.
- The solution efficiency perspective of the scorecard proved very difficult to implement and was not applicable in practice. During the use of the IDA method, the main obstacle was that data about resource consumption and the time needed for certain activities or the quality of activities "before" implementing the improvements detected either did not exist or were not made available due to confidentiality reasons. As a consequence, the indicators of the solution efficiency perspective were no longer captured. Instead, two new indicators were introduced: *perceived solution quality from customer perspective* and *perceived solution quality from method expert perspective*. Both were captured on a five-point scale.
- Many indicators needed refinements or adjustments. An example is *average learning time for new analyst until productivity*, where clarification was required as to whether self-study time also should be included in learning time and whether "productivity" means being able to contribute to the IDA method use or being able to use the method self-reliantly.

The indicator data collected with the scorecard were evaluated not only during InfoFLOW, but also during use of the IDA method in later years (see also Sect. 6.2). In every IDA use case, there were potentially four types of activities, which correspond to the phases of the IDA method: scoping, ID context analysis, demand modeling and consolidation. Every activity type potentially requires multiple steps (i.e., activities). For each activity, scorecard indicators were captured. For example, if demand modeling required several modeling sessions with different focus areas and participants, for each of the workshops indicators were captured as a separate activity.

The first group of indicators we selected for presentation in this paper originates from the method documentation quality and process efficiency perspectives of the scorecard. These four indicators were the ones preferred by the industrial partners in the InfoFLOW project who intended to use the method

for commercial purposes, that is, perceived productivity, perceived method value, perceived result quality (method user) and perceived result quality (client). All indicators used the same scale: 5—very good, 4—good, 3—acceptable, 2—improvements needed, 1—poor, 0—don't know. When preparing the data for presentation, we used two approaches:

- For all activity types in an IDA modeling case, we calculated the activity average for the case. Using these activity averages for a case, we calculated the overall average for a case. Indicator scores of "0" (don't know) were excluded in the average calculation. The case averages are shown in Fig. 7. The purpose of the chart was to visualize the general tendency of the method perception, here expressed in the four indicators, in order to check whether improvements made in the method handbook or the training material had any visible effect.
- The activity-type averages per case are shown in Fig. 8. Here, the intention was to see differences between activity types and to state where improvements should have priority.

Figures 7 and 8 are based on the same cases. Cases 1–5 were performed by method developers; Cases 6–10 were performed by other method users. After Cases 2 and 6, a new handbook version was released. Figures 7 and 8 are meant to illustrate indicator use in InfoFLOW-2. They are not meant to prove any statistically significant developments or correlations.

The indicator development shows improvements for *perceived method value* after Cases 2 and 6 when new handbook versions were released. *Perceived productivity* seems to be correlated to *perceived method value*, which is not surprising. When the method was used by its developers (Cases 1–5), the perceived result quality of the client was higher than that of method users. When the method was later used by other method users, it was the opposite. This indicates that method developers are more critical of the results or have higher expectations.

One of the main intentions referring to the activity-type averages was to detect which phase should have priority when working on improvements. In Cases 1 and 2, the scoping, demand modeling and consolidation needed improvement. With the new handbook published after Case 2, many of the problems were addressed. In demand modeling, to take one example, notation for the demand model was included that was previously missing. Cases 5 and 6 represent the phase of transferring the method knowledge from the method developer to the method user. Case 5 was done in cooperation between developer and user; Case 6 completely by a method user. The experiences from these first "external" uses resulted in the improvement of the handbook; that is, from Case 7, the new version was applied, which is also reflected in improved
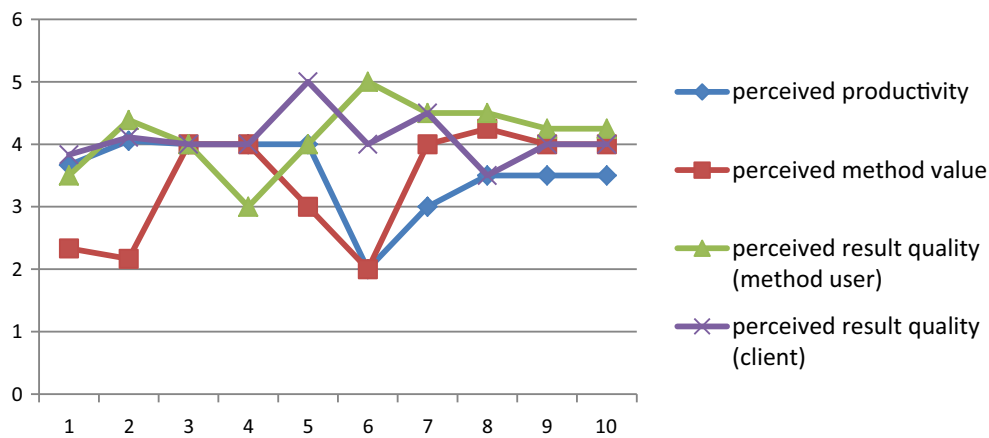
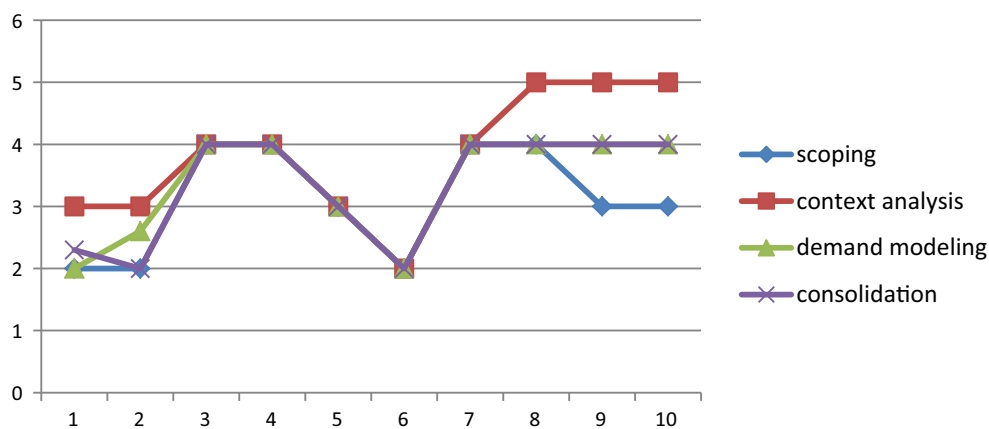**Fig. 7** Case averages of selected indicators

Fig. 8 Activity-type averages

activity-type averages. Currently, scoping seems to be in most need of improvement.

In the process efficiency perspective, we collected several indicators with the intention to detect potential needs for additional training of method users. The main indicators in this context were the number of times the method user already had performed the analysis task before, the time required for completing the analysis task and the perceived usefulness of the method support each time the analysis task was done. The expectation was that method users who needed much more time than the average or showed decreasing satisfaction with method support should be interviewed and maybe offered additional training. The experience showed, however, that if this indicator tendency appeared, the conclusions to draw from this were different from our assumptions.

Four method users showed decreasing satisfaction with method support during a period of 3 months. Three of these four stated that their satisfaction was decreasing because the more experience they got in the analysis task, the more they were looking for support on very specific modeling challenges occurring in their individual modeling cases that were

not covered in the method documentation. This did of course not indicate training needs, but rather potential areas for method extension or improvement. However, the actual support needs stated by the three method users were completely diverse and so specific to the actual modeling case that none of them was implemented as a method extension.

The fourth method user with this decreasing satisfaction tendency stated that at first he received sufficient support from the method documentation, but in the next steps was not sure how to apply it and then lost confidence in the process. This actually pointed to both the need for more training and method improvement. The method improvement part was taken care of in one of the new handbook releases.

With respect to the indicator value *time needed for the analysis task*, it quickly became clear that this was heavily dependent on the actual modeling case. For example, if the company under consideration had a quality handbook with clearly defined role definitions, the analysis of roles and demands was much faster than in companies without clear role definitions and a need to understand the tasks of different people first.

Table 8 Approaches inspiring indicator selection

Scorecard perspective	Approaches inspiring or providing indicators
Method documentation quality	Document quality indicators and document comprehensibility criteria [5], [44] and [45]
Pattern quality	Pattern quality from ontology engineering and enterprise modeling; conceptual model quality [47], [48] and [53]
Process efficiency	Business process management, task pattern use and method evaluation [47], [49], [50] and [51]
Solution efficiency	Effects of IT on organizational transformation [52]

6.2 Scorecard use by project-external IDA method users

In total, six different persons were trained in the IDA method and also used the scorecard in their IDA cases, which came from logistics, manufacturing, higher education and the IT industry. The scorecard indicators regarding learning time and perceived quality of the documentation were captured after the training. The other indicators were captured in every activity in each case (the same as in Sect. 6.1). Five different cases were the basis for this paper. Section 6.1 already presents the case averages and activity-type averages. Regarding learning time, Table 9 shows the time invested in training the different method users, separated into lecture-like training, self-study, working on examples or coaching in real cases. The table makes clear that training was intensified for later cases, which probably improved the understanding of the IDA method. This might explain the improved indicator value when comparing, for example, Cases 6 and 10.

6.3 Scorecard transferability to other methods

In order to check whether the scorecard would be applicable and transferable to other methods than IDA, we selected the 4EM enterprise modeling method as an example. Three persons used the scorecard in enterprise modeling cases with the 4EM method. The main intention was to investigate what

parts of the scorecard can be used without any changes for 4EM and where adaptations need to be made. Not in scope was the comparison of IDA and 4EM based on the scorecard values.

Before the method scorecard could be used for 4EM, all perspectives, aspects and indicators were checked for suitability for 4EM:

- Method documentation quality perspective: the aspects of documentation quality and method maturity could remain unchanged. Method support for pattern use and method support for pattern extension were not suitable and were removed, since 4EM does not include the use of patterns.
- Pattern quality perspective was not used—4EM does not use patterns.
- Process efficiency perspective: All three aspects of analysis process, analysis result (solution) and delivery were kept. As "analysis process" uses criteria and indicators that capture effort and duration for the different IDA method phases, these criteria had to be adapted to the activities of 4EM modeling.
- Solution efficiency perspective was not used because of the experiences undergone in IDA method improvement (see Sect. 5.1).

All three 4EM modelers managed to collect data about method documentation quality and process efficiency, which

Table 9 Learning time for the method users

Method user (case #) training part	Student 1 (7)	Students 2 & 3 (6)	Student 4 (8)	Student 5 (9)	Student 6 (10)
Lecture/presentation	2	2	4	6	6
Self-study of handbook	4	2	4	8	8
Exercise/example	0	0	2	4	4
Coaching on case	2	0	4	4	4
Total	8	4	14	22	22

Table 10 Indicators recommended for method improvement

Indicator	Relevant for sub-goal(s)	How to capture
Average learning time for new analyst until productivity	Method documentation quality	Captured by method specialist during training sessions
Perceived quality of completeness, correctness, understandability etc. on a suitable scale (e.g., five-point Likert scale)		Captured after each method use with a standard questionnaire
Number of improvements in documentation proposed by method users		Captured by method manager
Number of shortcomings detected in use		Captured by method manager
Average time for performing use of the method in a case	Process efficiency	Captured by the manager of modeling projects
Average time for performing a specific phase or component of the method		
Average cost for performing use of the method in a case		

confirms the feasibility of using the method scorecard for 4EM. However, in future work it should be investigated whether additional scorecard perspectives tailored to 4EM should be included. An example could be a perspective directed to participative modeling, an essential feature of 4EM.

7 Summary and future work

In the context of a method development project with a focus on information demand analysis and on improving information flow in organizations, the paper presented the development process of a scorecard intended to support method improvement. The article also presented the perspectives, aspects and (excerpts of) criteria of the method scorecard and illustrated its use for the IDA method and its transfer to the 4EM method. Among the conclusions to be drawn from this work are two rather "obvious" ones:

- The feasibility of scorecard development and the use as support for method improvement were demonstrated. Scorecard development helped to identify which criteria and indicators were important from the organizational method users' perspective.
- The transfer of the scorecard from IDA to 4EM indicates that many aspects and criteria are transferable between methods, although criteria reflecting the method phases needed adaptation. More cases are required to confirm and refine this.

In this context, we also learned from scorecard use which indicators are recommendable for method improvement projects or within an organizational content of method use. These indicators are summarized in Table 10. They are recommendable because either they contribute to monitoring several sub-goals, they are relatively easy to capture or they are simply essential for at least one sub-goal.

The more "hidden" conclusions are related to the utility of a scorecard: What are the actual benefits of using the scorecard? Could we have reached the same effects without the scorecard (i.e., without collecting and evaluating data)? Our impression is that the answers to these questions depend on the number of method users and cases of method use. For a method used by many persons in many cases, a sufficiently big "sample" is produced and the data collected will help to identify elements of a method that might be candidates for improvement efforts. However, the scorecard indicators should not be considered as the "only source of truth"; that is, the scorecard should be taken as a complementary means besides experience reports from method users. Section 6.1 shows an example: The indicators point at the scoping phase as a candidate for improving the method. This should be a motivation to investigate the scoping phase, but it does not mean that this part of the method is really the cause for the indicator values—there might be other causes, like the qualification of the modelers for "scoping," or the measurement procedure for the indicators might be inadequate.

Furthermore, some criteria and indicators of the scorecard need further investigation regarding their usefulness. An example is the average time required for the different phases of the IDA method. This time is partly dependent on the modeler and the complexity of the case. But if there are

many projects and different modelers, the development tendency of the average values of this indicator can be relevant.

Our preliminary recommendations regarding the method scorecard can be summarized as follows:

- Use the scorecard only for methods with many users and cases.
- For indicators addressing the time or the effort required for certain activities, find a way to normalize the complexity of these different activities.
- Consider reducing the number of indicators, e.g., to five per perspective.
- Use tool support for capturing and evaluating indicators.
- Use the scorecard as a complementary means for method evaluation and improvement only. Very valuable information for improvement of methods usually comes from the method users.
- Indicators can help in method evolution management.

Future work will on the one hand consist of continued data collection regarding the IDA method, which will probably lead to further development of the scorecard, and further investigation of transferability of the scorecard to other methods. Furthermore, more work is needed to understand the number of method users and cases for which scorecard use is appropriate. Investigation is also needed into whether a scorecard designed for method improvement for organizational purposes can also be applied as an instrument in method engineering. This is, to a large extent, a question of generalizability of scorecard perspectives and indicators, that is, whether the scorecard for a specific organizational context is also (in total or in part) valid for the general use of the method.

References

1. Sandkuhl, K., Seigerroth, U.: Evaluation framework, Version 1.2, InfoFLOW-2 Deliverable 1. School of Engineering at Jönköping University (2010)
2. Krogstie, J.: Model-Based Development and Evolution of Information Systems: A Quality Approach. Springer, Berlin (2012)
3. Benkenstein, M., Fellmann, M., Leyer, M., Sandkuhl, K.: The value of enterprise modelling: towards a service-centric perspective. In: Horkoff, J., Jeusfeld, M.A., Persson, A. (eds.) PoEM, Lecture Notes in Business Information Processing, vol. 267, pp. 299–306. Springer, Berlin (2016)
4. Kaplan, R.S., Norton, D.P.: The Balanced Scorecard: Translating Strategy into Action. Harvard Business Press, Boston (1996)
5. Märgner, V., Abed, H.E.: Tools and metrics for document analysis systems evaluation. In: Tombre, K. (ed.) Handbook of Document Image Processing and Recognition, pp. 1011–1036, Springer, London (2014)
6. Seigerroth, U.: Enterprise Modeling and Enterprise Architecture: the constituents of transformation and alignment of Business and IT. Int. J. IT/Bus. Alignment Gov. (IJITBAG) **2**(1), 16–34 (2011)
7. Goldkuhl, G.: Socio-instrumental pragmatism: a theoretical synthesis for pragmatic conceptualisation in information systems. In: Proceedings of the 3rd International Conference on Action in Language, Organisations and Information Systems (ALOIS), pp. 148–165. University of Limerick, Ireland, ISBN: 1-874653-79-8 (2005)
8. Lind, M., Seigerroth, U., Forsgren, O., Hjalmarsson A.: Co-design as social constructive pragmatism. In: The Inaugural Meeting of the AIS Special Interest Group on Pragmatist IS Research (SIGPrag 2008) at International Conference on Information Systems (ICIS2008), France. Sprouts—http://sprouts.aisnet.org/8-49 (2008)
9. Goldkuhl, G., Lind, M., Seigerroth, U.: Method integration: the need for a learning perspective. IEE Proc. Softw. **145**(4), 113–118 (1998). **(special issue on Information System Methodologies)**
10. Ralyté, J., Backlund, P., Kühn, H., Jeusfeld, M.A.: Method chunks for interoperability. In: Embley, D.W., Olivé, A., Ram, S. (eds.) Conceptual Modeling—ER 2006. Lecture Notes in Computer Science, vol. 4215, pp. 339–353. Springer, Berlin (2006). https://doi.org/10.1007/11901181_26
11. Mirbel, I., Ralyté, J.: Situational method engineering: combining assembly-based and roadmap-driven approaches. Requir. Eng. **11**(1), 58–78 (2006). https://doi.org/10.1007/s00766-005-0019-0
12. Brinkkemper, S.: Method engineering: engineering of information systems development methods and tools. Inf. Softw. Technol. **38**(4), 275–280 (1996). https://doi.org/10.1016/0950-5849(95)01059-9
13. Avison, D.E., Fitzgerald, G.: Information Systems Development: Methodologies, Techniques, and Tools. McGraw-Hill Education, Maidenhead (1995)
14. Henderson-Sellers, B., Ralyté, J., Ågerfalk, P., Rossi, M.: Situational Method Engineering. Springer, Berlin (2014)
15. Goldkuhl, G.: The grounding of usable knowledge: an inquiry in the epistemology of action knowledge. CMTO Research Papers, No. 1999:03, Linköping University (1999)
16. Lincoln, Y.S., Guba, E.G.: Naturalistic Inquir (Vol. 75). Sage, Thousand Oaks (1985)
17. Siau, K., Rossi, M.: Evaluating techniques for system analysis and design modelling methods: a review and comparative analysis. Inf. Syst. J. **21**(3), 249–268 (2008)
18. Harmsen, A.F.: Situational method engineering. Doctoral dissertation, University of Twente (1997)
19. Mooney, J., Gurbaxani, V., Kraemer, K.: A process oriented framework for assessing the business value of information technology. ACM SIGMIS Database DATABASE Adv. Inf. Syst. **27**(2), 68–81 (1995)
20. DeLone, W., McLean, E.: Information system success: the quest for the dependent variable. Inf. Syst. Res. **3**(1), 60–95 (1992)
21. Parker, M., Benson, R.: Information Economics. Prentice-Hall, Englewood Cliffs (1998)
22. Lewis, W.E.: Software Testing and Continuous Quality Improvement. CRC Press, Boca Raton (2016)
23. Kan, S.H.: Metrics and Models in Software Quality Engineering. Addison-Wesley Longman, Boston (2002)
24. International Organization for Standardization: Guidelines for the justification and development of management system standards. International Standard ISO Guide 72, Geneva (2001)

25. Deming, W.E.: Out of the Crisis. Massachusetts Institute of Technology, Center for Advanced Engineering Study, Cambridge (1986)

26. Kitchenham, B.: Procedures for performing systematic reviews. Keele University Technical Report TR/SE-0401. Keele, UK. ISSN:1353-7776 (2004)

27. Sandkuhl, K., Seigerroth, U.: Balanced scorecard for method improvement: approach and experiences. In: Enterprise, Business-Process and Information Systems Modeling, LNBIP, vol. 287, pp. 204–219. Springer, Cham (2017)

28. Sivaji, A., Zuaan, S.S.: Website user experience (UX) testing tool development using Open Source Software (OSS). In: 2012 Southeast Asian Network of Ergonomics Societies Conference: Ergonomics Innovations Leveraging User Experience and Sustainability, SEANES 2012. IEEE Xplore (2012). https://doi.org/10.1109/seanes.2012.6299576

29. Staron, M., Hansson, J., Feldt, R., Meding, W., Henriksson, A., Nilsson, S., Höglund, C.: Measuring and visualizing code stability: a case study at three companies. In: Proceedings: Joint Conference of the 23rd International Workshop on Software Measurement and the 8th International Conference on Software Process and Product Measurement, IWSM-MENSURA 2013. IEEE Explore (2013). https://doi.org/10.1109/iwsm-mensura.2013.35

30. Chang, J., King, W.: Measuring the performance of information systems: a functional scorecard. J. Manag. Inf. Syst. **22**(1), 85–115 (2005). https://doi.org/10.1080/07421222.2003.11045833

31. Keyes, J.: Implementing the IT Balanced Scorecard: Aligning IT with Corporate Strategy. Auerbach Publications, New York (2005)

32. Subramanian, G., Jiang, J., Klein, G.: Software quality and IS project performance improvements from software development process maturity and IS implementation strategies. J. Syst. Softw. **80**, 616–627 (2007)

33. Parmenter, D.: Key Performance Indicators: Developing, Implementing, and Using Winning KPIs. Wiley, Chichester (2015)

34. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design science in information systems research. MIS Q. **28**(1), 75–105 (2004)

35. Susman, G.I., Evered, R.D.: An assessment of the scientific merits of action research. Adm. Sci. Q. **23**(4), 582–603 (1978)

36. Wieringa, R., Morali, A.: Technical Action Research as a Validation Method in Information Systems Design Science, LNCS 7286. Springer, Berlin (2012)

37. Sein, M.K., Henfridsson, O., Purao, S., Rossi, M., Lindgren, R.: Action design research. MIS Q. **35**(1), 37–56 (2011)

38. Lundqvist, M., Seigerroth, U.: InfoFlow Application Case: Experiences from Modelling Activities at Proton Finishing. School of Engineering at Jönköping University, Sweden (2008)

39. Lundqvist, M.: InfoFlow Application Case: Experiences from Modelling Activities at Kongsberg Automotive. School of Engineering at Jönköping University, Sweden (2008)

40. Lundqvist, M., Seigerroth, U., Stirna, J.: InfoFlow Application Case: Experiences from Modelling Activities at SYSteam Management. School of Engineering at Jönköping University, Sweden (2008)

41. Lundqvist, M., Sandkuhl, K., Seigerroth, U., Stirna, J.: Method requirements for information demand analysis. In: 2nd International Conference on Adaptive Business Systems (ABS 2008). International Journal Communications of SWIN (CoSWIN), ISSN 1757-4439 (2008)

42. Lundqvist, M., Sandkuhl, K., Seigerroth, U.: Modelling information demand in an enterprise context: method, notation and lessons learned. Int. J. Syst. Model. Des. **2**(3), 74–96 (2011)

43. Sandkuhl, K., Lundqvist, M., Seigerroth, U.: Information demand patterns—concept, structure, and examples. InfoFLOW Deliverable D3, version 1.2, Jönköping University (2009)

44. Langer, I., Schulz von Thun, F., Meffert, J., Tausch, R.: Merkmale der Verständlichkeit schriftlicher Informations- und Lehrtexte. Zeitschrift für experimentelle und angewandte Psychologie **2**, 269–286 (1973)

45. Arthur, J.D., Stevens, K.T.: Document quality indicators: a framework for assessing documentation adequacy. J. Softw. Evol. Process **4**(3), 129–142 (1992)

46. Sandkuhl, K.: Information demand patterns. In: Proceedings of PATTERNS 2011, the Third International Conferences on Pervasive Patterns and Applications, pp. 1–6. September 25–30, 2011, Rome (2011)

47. Sandkuhl, K.: Capturing product development knowledge with task patterns: evaluation of economic effects. Q. J. Control Cybern. **39**(1), 259–268 (2010)

48. Moody, D., Shanks, G.: What makes a good data model? Evaluating the quality of entity relationship models. In: Entity-Relationship Approach|ER'94 Business Modelling and Re-Engineering, Springer LNCS, vol. 881, pp. 94–111. Springer, Heidelberg (1994)

49. Aalst, W.M., Marcelo, L.R., Maria, S.: Business process management don't forget to improve the process. Bus. Inf. Syst. Eng. (2016). https://doi.org/10.1007/s12599-015-0409-x

50. Weske, M.: Business Process Management: Concepts, Languages, Architectures, 2nd edn. Springer, Berlin (2012)

51. Sandkuhl, K., Tellioglu, H., Johnsen, S.: Orchestrating economic, socio-technical and technical validation using visual modelling. In: ECIS 2008 Proceedings, pp. 1752–1763. AIS Electronic Library (AISeL) (2008)

52. Gregor, S., Martin, M., Fernandez, W., Stern, S., Vitale, M.: The transformational dimension in the realization of business value from information technology. J. Strateg. Inf. Syst. **15**(3), 249–270 (2006)

53. Hammar, K.: Towards an Ontology Design Pattern Quality Model. Linköping Studies in Science and Technology. Linköping University, Sweden (2013)

Kurt Sandkuhl is a professor of Business Information Systems at the Institute of Computer Science, University of Rostock (Germany), and affiliate professor of Information Engineering at Jönköping University (Sweden). He received a diploma (Dipl.-Inform.) and a Ph.D. (Dr.-Ing.) in Computer Science from Berlin University of Technology in 1988 and 1994, respectively. Furthermore, he received the Swedish degree as "Docent" (postdoctoral lecturing qualification) from Linköping University, Institute of Technology, in 2005. His research interests include the fields of enterprise modeling, ontology engineering and method engineering. Kurt Sandkuhl is author or co-author of five books and more than 200 peer-reviewed publications.

Dr. Ulf Seigerroth is professor of Informatics at Jönköping University (JU), School of Engineering (JTH). Seigerroth holds a Ph.D. in Information System Development from Linköping University (2003). In 2004, he was appointed head of the Informatics Department at Jönköping International Business School (JIBS), which was a position that he had for almost 4 years. During this time, Seigerroth was the co-founder and co-director of CenIT (Centre of Evolving IT in Networked organizations) and he was also the co-founder of the graduate school of informatics (GSI) at JIBS. In 2011, he was appointed associate professor of Informatics (docent) at JTH. During 2011–2017, Ulf was director of the research environment in Computer Science and Informatics at JTH. Since 2016, Ulf is a member of the Management Board for the KK-environment at JU, Knowledge-Intensive Product Realization (SPARK). During his employment at JU, Ulf has a strong engagement in education on basic, advanced and Ph.D. level. Ulf has also been program manager for the bachelor program in Business Informatics. After his Ph.D., Ulf was also program manager for two master programs at JIBS.