Effective and Efficient Information Quality Assessment

1 Scientific Questions and Industry Problems

Technical documentation, such as user manuals and technical specifications, often constitutes the first line of support when users need help or want to learn more about a product or a service. It becomes an important part of the user experience and the information provided needs to be of appropriate quality, e.g. correct and relevant. If technical documentation is inadequate, or even wrong, the consequences can be serious.

In 2000, 28 cancer patients were exposed to radiation overdoses due to defects in the software that controlled a radiation machine and it's manual [8]. The doctors and physicians that treated the patients wanted to increase the number of shield blocks used to protect sensitive organs from the harmful radiation from four to five, which was possible according to the manual. However, the software could only reliably handle up to four shield blocks, and in 28 out of 56 cases, patients suffered doses that were 65% too high on average. As a result, at least 17 patients died from radiation poisoning.

In 2004, an F/A-22 aircraft crashed on initial takeoff from the runway [17]. The pilot ejected safely and sustained only minor injuries. The aircraft was completely destroyed, also causing damages to the runway. The loss was valued at more than $133 million. The cause of this mishap was a power interruption, which made the aircraft uncontrollable. The pilot was unaware of this condition because he did not perform a test prior to takeoff because of ambiguous technical orders which lead to an incorrect understanding of the situation. The Accident Investigation confirmed that this understanding was widespread among the personnel and based on poor technical data system descriptions.

While there are several issues involved in both cases, a lack of *Information Quality* (IQ) of the technical documentation is a significant factor. Methods and means to assess information quality have been the focus of much research. A major difficulty is that quality is a multidimensional and often subjective concept. We use it in daily language as an intangible trait, something that can be felt or judged, but often not exactly measured or weighted. In order to manage quality, to be able to assess and improve it, it has to be formally defined and objectively measured. The **purpose** of the proposed project is to *improve the effectiveness and the efficiency with which IQ can be assessed*. We are guided by the following **scientific questions**:

SQ1 What is the necessary theoretical foundation for IQ assessment, with respect to definitions, models and methods?

SQ2 What processes and tools can be used to assure effective and efficient IQ assessment?

SQ3 Is there a theory that unifies quality assurance of software and information?

Research questions SQ1 and SQ2 are highly relevant for our industry partners, and the information engineering society in general. If we can answer these questions in a positive and constructive way, we can use the results to address the following **industry problems**:

IP1 How can we best use effective and efficient technology for IQ assessment in a production environment?

IP2 Can we quantify and predict the benefits of applying this technology in terms of reduced overall costs and/or increased customer satisfaction?

Both owners of technical documentation as well as companies that offer information engineering tools and services will benefit from solutions to these industry problems. Owners, such as Ericsson, can provide increased customer satisfaction at a lower cost. Service and technology providers, such as Sigma Kudos, can provide more effective and efficient tools and services, and better highlight the benefit of these.

Research question SQ3 goes beyond information quality and will result in new insights into software and system engineering. A positive and constructive answer can result in methods for a holistic approach to quality assessment of software (code, formal specifications, etc.) and information (manuals, documentation, requirements specifications, etc.). In the general case, such an approach can result in technical products and services of higher quality, and in turn lower costs and better user experiences. In extreme cases, it can reduce the risk of financial losses, injuries, or loss of human lives.

2 Scientific and Industry Goals, Operative Results

We can define a number of **scientific goals** to help provide an appropriate theoretical foundation and technology to answer scientific questions SQ1–SQ3. These scientific goals are:

SG1a Develop suitable definitions of IQ.

SG1b Define and validate IQ models adaptable to different users and usage scenarios of technical documentation that allows for effective and efficient assessment based on the definitions from SG1a.

SG1c Define and validate methods to adapt the IQ models from SG1b to specific users and usage scenarios of technical documentation.

SG2a Define and validate processes and process support to implement the methods from SG1c.

SG2b Design, implement, verify, and validate generic tools for effective and efficient IQ assessment based on arbitrary IQ models that were developed in SG1b.

SG3 Develop a theory and experimental tools that can be used to assess the quality of a software system by considering the quality of both software and information.

The scientific goals correspond to the scientific questions. The theory developed to address the scientific questions will form the basis for the tools and methods developed. Similarly, the tools and methods developed will serve as proof of concept of the theory.

We define a number of **industry goals** that show the practical applicability of the scientific results and demonstrate the business value of our approach compared to state-of-the-art methods:

IG1a Show that the processes and process support tools from SG2a can be used to design quality models suitable for specific users and usage scenarios of technical documentation (e.g., that owned by Ericsson).

IG1b Show that the generic tools from SG2b can be used to assess IQ in a production setting.

IG2a Evaluate the costs and benefits of the suggested IQ model (IG1a) and assessment (IG1b), and quantitatively compare these to industrial practices in use today (e.g. at Ericsson).

IG2b Predict costs and benefits from using IQ models (SG1a) and assessment (SG1b) for new users and usage scenarios to allow service and technology providers (e.g. Sigma Kudos) to quantitatively compare them to the state-of-the-art methods of IQ assessment.

The industry goals will leverage the results of the science goals, specifically the tools and methods. These will be integrated with current practice and technology in use. For example, IG1a and IG1b will be accomplished by extending DocFactory, a real-world Content Management System (CMS) provided by Sigma Kudos, with the results of scientific goals SG1 to SG2. The new system will be used to assess technical documentation from Ericsson, to investigate how well the technology fits an industrial setting (IP1).

IG2 is achieved by correlating IQ assessed with both our technology and state-of-the-art methods, and by developing a prediction model for costs and benefits of IQ assessment. This not only solves the industry problem of quantifying the benefits of our approach, but also serves as input to scientific experiments that are conducted to validate the effectiveness and efficiency of our approach.

3 Project Description

3.1 State of the Art

General quality assessment and management processes and practices are well established and standardized, e.g. in the ISO 9000 family of standards [13].

In order to discuss and assess the quality of information, we first need to define it. Crosby [7] defines quality as *"conformance to requirements"* suggesting that there exists a set of well-defined requirements. Another notion of quality is given by Juran [14] who defines it as *"fitness for use"*. This definition considers the customers, and their demands and expectations. To define IQ, we need both notions. On one hand, the demands and expectations of the customers guide requirements of a technical documentation and conformance to well-motivated requirements contributes its fitness for use. On the other hand, not all diverse and shifting user expectations can be boiled down to requirements. There exist several elaborations of and quality frameworks for IQ, e.g. by Ge and Helfert [9], Kahn et al. [15], Klein et al. [16], Wang and Strong [24]. However, all related IQ assessment methods are subjective and manual, e.g. based on expert knowledge or checklists. While these methods can be effective, they are inefficient.

In the related field of software quality assessment, McCall et al. [22] present a *software quality model* defining quality factors corresponding to the stages of a software life-cycle and relating metrics that are used to measure these quality factors. Several metrics are weighted and used to determine the quality of each factor. Many metrics are based on checklists, which means that they are subjectively measured. McConnell [23] differentiates between internal and external quality, i.e. quality that affects the product while produced and maintained vs. quality when the product is in use. All this is standardized in the ISO/IEC 9126 family of standards [12]. The model by McCall et al. introduces several important ideas. First, there is not *one* software quality, but several factors that affect quality. Second, these factors matter during different periods of the life cycle. Third, the quality factors should be measurable and metrics should be defined. Several modern software quality models and metrics suites exist, such as those by Li and Henry [18], Chidamber and Kemerer [6], Abreu [1], and Henderson-Sellers [11]. There are several studies that validate the claim that metrics can be used as an effective indicator of the quality of software, for example Basili et al. [5] and Harrison et al. [10]. However, our work on metrics validation shows that existing software metrics tools and software quality models give contradicting advises and should be used with care [19, 20].

Our contributions combine IQ definitions and frameworks with metrics and methods from software quality as well as software visualization to help automate IQ assessment [25, 26, 27, 29, 30]. We also apply analyses such as coverage to support manual IQ testing. We have adapted the VizzAnalyzer[1], a framework for software analysis and visualization [21], to IQ analysis and visualization [28]. This is to best of our knowledge the first time automated methods have been

[1] `http://www.arisa.se/vizz_analyzer.php`

used to measure, test, and visualize IQ.

3.2 Scientific Method and Approach

A significant part of the proposed project is focused on theory development. The theory will be implemented by methods and tools, which will be validated in the practical information engineering projects together with the industry partners.

The project will use a spiral model, where theoretical insights drive method and tool development. Each innovation cycle starts from a set of ideas and validated solutions. It results in a set of issues, questions, and a plan on how to address these. The cycles follow the traditional scientific method, where a model is created and used to formulate hypotheses. The hypotheses are validated through experiments, which in turn are used to refine the model. Each cycle pushes the knowledge boundary towards the envisioned goals. In the proposed project, we expect to complete two such innovation cycles. In each of them, our scientific approach is to:

1. Adapt software quality models based on measurement and testing to IQ assessment.
2. Experimentally assess IQ of real-world technical documentation with our *engineering approach* based on measurement and testing as well as with the manual *state-of-the-art approaches* (cf. Section 3.1).
3. Validate the effectiveness by correlating assessment results of the engineering and the state-of-the-art approaches.
4. Validate the efficiency by comparing time and resource requirements of the engineering and the state-of-the-art approaches.
5. Abstract the findings in IQ assessment together with the earlier findings on software quality assessment to develop a theory of quality assessment uniformly and holistically applicable to software and information artifacts of (software) systems.

We use the Goal-Question-Metric (GQM) approach [4] to establish a generic IQ model. We define a set of possible quality *goals* selected and weighted depending on users and usage scenarios of technical documentation. These quality goals are backed by *questions* that aim to determine if the goals are fulfilled or not. Each question is answered using one or more *metrics*, referred to as *Key Performance Indicators* (KPI). For example, if understandability is a goal, a suitable question tests readability, which in turn can be assessed by KPIs like orthographical and grammatical correctness and appropriateness of cross-references. Today, industrial KPIs are used as manually assessed IQ benchmarks. Our pre-studies show that many KPIs can be assessed using automatic *measurements*. They further show that many relevant KPIs that cannot be assessed automatically—and that are not correlated to such—can be assessed by *testing*. Users of the technical documentation are asked to perform certain tasks and the outcome and their use of the documentation is monitored. The IQ testing can be additionally supported by indirect metrics, such as the coverage of the technical documentation.

We rely on statistical analyses to validate how *effective* the quality assessment methods are. We assume that there exist a set of well-defined and well-proven (effective, but inefficient) state-of-the-art methods to assess KPIs, for example expert knowledge and checklists. Our engineering approach assesses them using IQ measurement and testing, adapted from software measurement and testing. Our validation experiments correlates results from IQ measurement and testing against the outcomes of the state-of-the-art IQ assessment. The basis for the statistical analysis of such experiments is hypothesis testing [2]. A null hypothesis H_0 is defined formally. The data collected during the course of the experiment is used, if possible, to reject H_0 to draw the opposite conclusion. In our concrete case, H_0 states that correlations of the two approaches of assessing IQ—engineering, using the IQ measurement- and testing, and state-of-the-art, using checklists—are only coincidental. The null hypothesis should be rejected with as high significance as possible. We start for all our analyses with the standard borderline significance level of 0.05, i.e. observations are not coincidental but significant with at most a 5% error possibility (significance value $p < 0.05$). If we can reject the null hypothesis, the alternative hypothesis H_1, namely that the state-of-the-art approach is as accurate as our engineering approach, can be assumed.

In this case, the engineering approach has *efficiency* advantages over the state-of-the-art approach: the engineering approach is (semi-)automated and objective, while the state-of-the-art approach is based on human judgement of KPIs and, hence, less efficient and subjective.

The four phases specified by Basili et al. [3] describing the experimentation process in software engineering form the basis of our project implementation and the layout of our work packages (cf. Section 3.4). The first phase is the definition phase, i.e. deciding on motive, objective, purpose, etc. (WP 1). The second phase is the planning phase, i.e. setting up the experiments (WPs 2-5). The third phase is the operation phase, i.e. conducting the experiments (WP 6). The fourth phase is the interpretation phase, i.e. analysis and putting the statistics in broadening series of context (WPs 7-8).

3.3 Competence and Contributions to the Project

The proposed project requires three major areas of expertise: software and information quality analysis, information engineering of technical documentation, and users and usage scenarios of technical documentation. The three partners together provide expertise in all the required areas. We demonstrated this in our pre-studies, where Software Technology Group and Sigma Kudos jointly applied software measurements and testing to technical documentation provided by Ericsson and others. The results were presented on an open seminar series on "The future of technical information"[2], organized by Sigma Kudos.

The Software Technology Group (STG) at Linnaeus University has established itself internationally as one of the leading groups in the field of software and information quality analysis. In collaboration with industry, the group has contributed to automated assessment and validation of software quality metrics in several ways, including a project funded by the Knowledge foundation[3]. The project was evaluated as an "Interesting and good project which led to both scientific publications and co-production with industry and even a spin-off company."

The group was recently awarded a VINNOVA VINNMER grant[4], where the research on information quality and the research milieu were considered highly important and of excellent quality.

Prof Dr Welf Löwe, Dr Morgan Ericsson, and Dr Anna Wingkvist have jointly developed automated IQ assessment as a scientific field. In less than two years, we have published our research results in several conference publications and two invited journal articles. Welf has a background in software analysis and has published more than 100 scientific articles in the field; his most influential article on software pattern detection has more than 120 citations. He is also an experienced project manager. Morgan and Anna have, in less than three and two years after completing the PhDs, respectively, acquired significant background knowledge and research experience in information systems development, integration, and evaluation.

The STG maintains and develops several open source analysis tools, such as the VizzAnalyzer. It has been actively developed and improved by the group for more than 10 years, and numerous visualizations, and software and information analyses have been integrated since.

---

[2]The seminars took place in Göteborg, Nov. 2009; Malmö, Feb. 2010; Stockholm, Apr. 2010.

[3]PI Welf Löwe with Alstom, Artisan, Combitech, and Windh, DNR: 2005/0218, Grant: 1,735 MSEK

[4]PI Anna Wingkvist, DNR: 2011-01351, Grant: 2,1 MSEK

WPs	Time	Work-load	Anna (STG)	Morgan (STG)	Welf (STG)	Mats/Liselotte (Ericsson)	Johan (Sigma Kudos)
WP1	6 months	9	1	3	1	2	2
WP2	2 months	3	2	0.25	0.25	0.25	0.25
WP3	6 months	9	0.25	1.5	5	0.25	2
WP4	6 months	9	0.5	6	0.25	0.25	2
WP5	4 months	6	2	1.5	0.25	2	0.25
WP6	4 months	6	2	1.5	0.25	2	0.25
WP7	2 months	3	0.25	1.25	1	0.25	0.25
WP8	6 months	9	1	3	1	2	2
Sum	36 months	54	9	18	9	9	9

Table 1: Work package total time ("Time") and workload, total ("Workload") and distributed over project participants. The figures are approximations, and will be adapted if needed. They are given in person months with the exception of "Time", which is given in months.

To promote and support industry collaborations, the STG established the Information Engineering Center (IEC, `lnu.se/IEC`) — a non-commercial cluster that currently has 60 partners from industry, academia, and the public sector jointly working with knowledge transfer, education, and research in the field of information engineering. Welf is the chairman and Sigma Kudos was one of the first members.

Sigma Kudos manages technical documentation for companies and provides information engineering expertise. DocFactory, their CMS, is used to create, maintain and distribute a variety of technical documentation. Their participation in the project will provide insights into the processes and software used to produce technical documentation, as well as the requirements during the process. Sigma Kudos supplies the project with licenses to DocFactory, worth 1,44 MSEK, as well as access to real-world technical documentation and requirements from their customers. Johan Thornadtsson is an expert on DocFactory and has considerable experience of how Sigma Kudos works with technical documentation. He deals especially with Ericsson, which is one of Sigma Kudos' customers. He will contribute approximately 9 person months (about 1,500 hours).

Ericsson is a customer of Sigma Kudos and uses DocFactory for their technical documentation. They are focused on quality improvement of their technical documentation. Ericsson will provide real-world technical documentation for two systems, as well as expertise on users and usage scenarios of these two documentations. Mats Slunga is a systems manager responsible for the Gateway GPRS Support Node (GGSN), a major component of the GPRS network. Liselotte Wanhov is a systems manager responsible for the Converged Packet Gateway (CPG). They will provide the customer perspective of and insights into technical documentation engineering and quality management integrated in a larger system and engineering context. Mats and Liselotte will together contribute approximately 9 person months (about 1,500 hours).

3.4 Work Packages

The project is divided into Work Packages (WPs), that correspond to the overall scientific approach. Here we define the purpose, milestones, responsible persons of each WP, and give a brief description. The duration and workload of the WPs is summarized in Table 1 and their dependencies are depicted by Figure 1.

Anna/Morgan/Welf (STG) are the principal investigators and responsible for most WPs. Johan (Sigma Kudos) and Mats/Liselotte (Ericsson), will provide CMS tools, technical documentation,

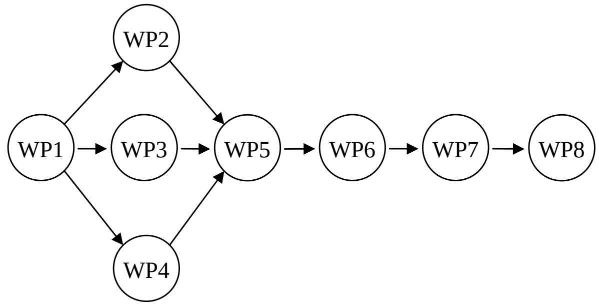

Figure 1: Dependencies between work packages (will be scheduled over two innovation cycles).

and user expertise. They are responsible for some WPs and involved in others to some degree.

WP 1 — IQ Modelling and Theoretical Validation

Purpose: Define an IQ model that supports (semi-)automated assessment. **Milestone:** A compendium with a quality model and a list of KPIs as requested by Ericsson and Sigma Kudos. KPIs are mapped to metrics and ranked in order of importance. **Responsible:** Morgan (STG).

The first step is to create an understanding of the concept of IQ and our different views on it. Based on this common understanding, we will use the GQM approach to establish a quality model including KPIs. This quality model will be inspired by, but no limited to, experiences from the preliminary case studies. Therefore, we will organize two workshops on quality. The first one will introduce how we deal with software quality, information quality, etc., both on a theoretical level as well as practical. Results of the preliminary studies will be presented. Concepts such as GQM will be introduced, and tools like the DocFactory and VizzAnalyzer will be demonstrated. The workshop will end with a short brainstorming session, where we attempt to establish common quality goals. The second workshop will focus on defining the quality model, and it contains several brainstorming sessions and discussions.

Once the common quality model is established, we focus on the identified metrics supporting KPIs. The aim is to find suitable metrics, e.g. software metrics that can be adapted and/or reused. The metrics are divided into two categories: (a) Metrics that can be assessed automatically, and (b) Metrics that require some manual intervention, i.e. the metrics are only in part supported by automation. Metrics of category (a) will assess structure, and the measurements will be conducted on the level of a descriptive markup language (i.e. XML, HTML, or similar). The metrics of category (b) will focus on qualities that in part require a subjective stance, e.g. reliability, suitability, understandability, etc. Depending on their number, the metrics will also be ranked and ordered according to priority. Results will be discussed in a workshop before being added to a compendium.

We rely on our scientific expertise in software and IQ assessment (Anna, Morgan, Welf) and the industry expertise in technical documentation production and usage (Johan, Mats, Liselotte) to perform the categorization and ranking. Together, we validate the applicability of the quality model in theory and practice. Especially, Sigma Kudos validates the computability of metrics based on the information and meta-information captured in DocFactory, while Ericsson validates the applicability of metrics for the two cases: the GGSN-MPG and CPG CPIs.

WP 2 — Visualize Metrics Data

Purpose: Define visualizations to communicate quality. **Milestone:** A list of visualizations, ranked in order of importance, implemented and integrated in the tools. **Responsible:** Anna (STG).

Metrics, assessed directly or indirectly via testing, are an important part of the quality model, but the ability to communicate and visualize results is just as important, e.g. to discuss their ranking or their actual relation to goals in concrete case studies. In fact, visualizations allow a theoretical pre-validation of measurement and testing using human experts.

We will study a number of visualizations known from software and information engineering, and decide which of these can be adapted to IQ. Our industry partners will play an important role, since visualizations can be integrated into IQ management. They will offer feedback on the understandability and usability of the visualizations.

The visualizations will be implemented in VizzAnalyzer based on information and meta-information captured in DocFactory. They will be discussed at a workshop targeting views on technical documentation quality issues.

WP 3 — IQ Measurement

Purpose: Methods and tools for IQ Measurement. **Milestone:** Implements metrics for IQ assessment from category (a) using VizzAnalyzer and information exposed by DocFactory. The tool extensions are tested and performance-evaluated on real-world technical documentation. A journal publication is submitted. **Responsible:** Welf (STG).

The IQ metrics of category (a) need to be defined exactly, theoretically evaluated, implemented, and tested. To formally define them, we first need to formalize information extracted from technical documentation, i.e. define a suitable meta-model, which will be based on tree grammars and relational algebra. The metrics will be defined using this meta-model.

We will implement the metrics using VizzAnalyzer and integrate them with DocFactory. If DocFactory exposes information of technical documentation as specified in the meta-model, we can show that the implemented metrics conform to their algebraic specifications.

The implementation will be the managed by Welf. The industry partners will supply DocFactory, test data, and support the tool integration. More specifically, Johan will provide information required to adapt DocFactory's information and meta-information export if necessary. He will also support the integration with VizzAnalyzer's structural analyses and their visualizations. Mats/Liselotte will provide cases to test the metrics implementations, and the changes made to DocFactory and VizzAnalyzer. They will also provide feedback.

WP 4 — Information Testing

Purpose: Process and tool support for information testing. **Milestone:** A process for information testing. Logging facilities and metrics that gather data from the logs are implemented, tested and performance-evaluated on real documentations. A journal publication is submitted. **Responsible:** Morgan (STG).

The so-called "-ilities" will require some manual "testing". We focus on combining automated metrics and procedures from software testing to define processes to test aspects of IQ using metrics that fall into category (b).

The information testing will be conducted by manual procedures supported by automated analyses. The process to carry out the tests needs to be carefully described. We establish this process and decide what artifacts, such as test suits, test cases, assertions, etc., that need to be produced. We will also adapt the notions of test coverage and test ordering to information. Finally, the test results of individuals testers need to be aggregated to a metrics value in order to be able to integrate them with metrics from category (a). So, suitable statistical methods need to be adopted and theoretically validated.

We will develop an information testing environment based on DocFactory as an information *browser*. It will functions like a regular Web browser or PDF reader, but extended with logging facilities that support coverage analyses. The logging information will be fused with structural information, i.e. corresponding analyses need to be integrated in VizzAnalyzer. Morgan will manage the implementation and Johan will provide the necessary documentation and access to DocFactory.

WP 5 — Experimental Setup
Purpose: Plan the experiment. **Milestone:** Detailed description of the experiments, seamlessly integrated tools. **Responsible:** Anna (STG) and Mats/Liselotte (Ericsson).

The models implemented in WP 1 and the tools and processes developed during WP 3 and WP 4 will be validated by experiments. We plan to conduct a series of realistic experiments using real-world technical documentation and test users to achieve this. The experiments will be conducted in collaboration with Sigma Kudos and Ericsson. Mats/Liselotte will provide real-world technical documentation, test users, and a test suite containing test cases with expected user behavior. They also serve as human experts and judge certain qualities in their respective documentations.

A workshop will define the set-up of the validation experiment. This validation experiment will correlate results from automated measurements and (semi-)automated testing – both integrated using our IQ quality model from WP 1 – with results of expert-based IQ assessment.

WP 6 — Experimenting
Purpose: Gain statistical data for validation and experience to help determining the business value. **Milestone:** Detailed description of the experiments, how they were performed including time and effort, and the data gathered. Deviation from the plans of WP 5 will be documented and motivated. **Responsible:** Anna (STG) and Mats/Liselotte (Ericsson).

The experiment planned in WP 5 will be carried out in collaboration with our industry partners. More specifically, Mats and Liselotte will supervise information testing of GGSN CPI and CGP CPI, respectively. Mats and Liselotte will also provide expert assessments of their respective technical documentation using state-of-the-art approaches. Johan will apply IQ measurement and collect data from other customers of Sigma Kudos. As an expert, he will judge the qualities of these technical documentations. Fully automated measurements of metrics (a) will be applied to a variety of technical documentations from different customers of Sigma Kudos. Semi-automated measurements of metrics (b) collecting user behavior data will only be applied to Ericsson's documentations because of their required (manual) effort.

WP 7 — Evaluation of Experiment
Purpose: Validate of our approach and correlate our engineering approach and traditional approaches. **Milestone:** Hypothesis H_0 rejected. **Responsible:** Welf (STG).

In order to validate the effectiveness of our engineering approach to IQ assessment, based on measurement and testing, we rely on statistical analysis. We formally define a hypothesis and plan experiments (WP 5) and collect data during the course of several experiments (WP 6). This data is used to reject the H_0 hypothesis and, if this is possible, to draw the desired conclusions, cf. Section 3.2. Welf is responsible for this analysis.

In order to determine the business value of IQ measurement and testing, we rely on benchmarks and experiences that the industry partners gained during the the experiments. Johan, Mats and Liselotte are responsible for assessing the exact efforts and benefits of the engineering and the state-of-the-art approach and to project it to other cases. More specifically, Johan will determine the effort and effect of assessing IQ using state-of-the-art approach. Mats and Liselotte will determine the effort and effect of IQ assessment using their current methods compared to our information testing approach. From this assessment, Welf will develop an initial prediction model based on this data and discussions with Johan, Mats, and Liselotte.

WP 8 — Theory Building and Conclusions
Purpose: Develop a uniform theory around software and information quality assessment and dissemination of the results. **Milestone:** A better understanding of scientific questions SQ1–SQ3 and industry problems IP1, IP2. A journal publication is submitted. **Responsible:** Morgan (STG).

Conclusions: We will have gained and validated the theoretical results of WP 1 to address SQ1. The results of WP 3 and WP 4 form the basis for an evaluation of the efficiency of our approach. We combined this with a validation of effectiveness, a result of WP 7, to address SQ2. The results will be summarized at a workshop before preparing the final journal publication and an updated compendium.

We will use the outcome of WP 6 and 7 to summarize the benefits and drawbacks of our engineering approach to IQ assessment compared to the current state-of-the-art approaches. We address IP1 in the context of Ericsson and other customers of Sigma Kudos, and provide an initial prediction model to address IQ2.

Theory building: The commonalities and differences compared to software measurement and testing will serve as input to further theory building. We will compare the set of goals, questions, and metrics (including their assessment by measurement and testing) as proposed for software and information and factor out commonalities. We will analyze semi-formal information artifacts with software and information quality assessment technology. Finally, we will identify references (explicit or implicit ones) between formal and informal information artifacts. The outcome of SQ3 will be considered a scientific hypothesis, which will require experimental validation. We plan to address this in a future project.

3.5 Expected Results and their Significance for Science and Industry

Scientific Results We expect the proposed project to produce two main research results. The first is an effective and efficient approach to IQ assessment, that is scientifically validated (and tested in practice). Our approach contains suitable IQ definitions (SG1a), appropriate generic quality models (SG1b) and method and tools for their adaptation to users and usage scenarios of technical documentation (SG1b, SG2a), tools for IQ measurement, testing, and communication (SG2b). Altogether, this result answers research questions SQ1 and SQ2 constructively and improves the state-of-the-art IQ management.

The second result is an integrated and holistic theory of effective and efficient quality assessment of products and services, the software contained, and the technical documentation describing them (SG3/SQ3). Every information artifacts connected to a product or a service — informal text like descriptions of requirements and user manuals, semi-formal like use-case and requirement specifications, or formal like architecture and design specifications and actual implementation — should conform to another as they describe the *same* product or service from different perspectives and at different levels of abstraction. However, the state-of-the-art approach assesses the conformance of these different artifacts at best manually and subjectively. The quality of these artifacts is often only assessed in isolation (e.g. proof-reading, unit testing). A holistic view on quality of these artifacts that overcomes the cognitive boundaries when assessing informal, semi-formal and formal artifacts, helps remove inconsistencies and improve quality of technical products and services in general.

The results of the proposed project will be published in scientific journals, and presented at conferences and workshops. It will provide Anna and Morgan with research time, which will help them further their careers and qualify as Docents at the (successful) completion of the project.

The proposed project will have an impact on courses on "Software Quality" and "Applied Program Analysis" which will evolve towards courses on "Software and Information Quality" and "Applied Program and Information Analysis", respectively. The research partners are responsible for the course development and industry partners are invited to contribute and participate in lectures. Moreover, the project will create a number of Master's Thesis projects co-supervised by research and industry partners. These projects will provide preliminary studies before deploying new tools and processes and evaluations of statistical results from different angles. The courses and master's projects will support technology transfer by exposing students to practical and up-to-date topics that make them employable by industry.

Results for Industry In our preliminary experiments, we applied certain software metrics to assess IQ based on Sigma Kudos' DocFactory and STG's VizzAnalyzer: the former captures technical documentation and logs its usage, the latter analyzes and visualizes the technical information and its usage. The two tools were only loosely coupled. This loose coupling, which required some adaptation for each concrete technical documentation, is insufficient to provide a statistical basis for the scientific validation required; a coupling with high performance and convenience will be one of the practical results contributing to the industry goal IG1b. It allows IQ monitoring while producing, maintaining, and using technical documentation without hampering the core activities. We will produce such a system by integrating DocFactory and VizzAnalyzer. Sigma Kudos gains access to means to work with IQ, both towards customers and internally during the production of technical documentation.

The GQM approach to IQ can be used to negotiate quality requirements and KPIs with customers, and the metrics and visualizations can provide means to automatically assess quality during the information engineering process contributing to IG1a.

Ericsson gains a better control of IQ and a way to reason about it. This contributes to goal IG2a: Objective means to assess IQ makes it possible to plan, calculate costs, and validate the final technical documentations. Ericsson will also gain access to a number of tools to specify and investigate the quality of technical documentations, before, during, and after production. These can be used to define and gradually improve prediction models for costs and benefits of IQ in information engineering processes (IG2b). These models can be used to plan (Ericsson) and to support sales (Sigma Kudos), for example.

We expect (semi-)automated IQ assessment to be seamlessly integrated in real-world CMS products and information engineering processes ready for use by any owner of technical documentation. The assessment-evaluation-feedback process — with a constant monitoring of production, maintenance, and usage of technical documentation — becomes manageable for the different stakeholders of a technical documentation. Experiments and practical experiences in the project based on several real-world information engineering projects show effectiveness and efficiency of this approach and demonstrate its business value. It addresses IP1 and IP2 for the information engineering industry, in general.

We will organize public seminars for IEC members and other interested parties that demonstrates the opportunities of (semi-)automated IQ assessment. We will also define an IQ compendium including IQ models, KPIs, metrics, and processes for assessing IQ. We have had good experiences with a similar compendium on software quality metrics and models[5] in our previous Knowledge Foundation project. This compendium is a value in itself as it defines our (current) notion of IQ as assessable by our approach. This compendium will be made public, available online as an interactive document.

[5] `http://arisa.se/compendium/`

References

[1] F. B. Abreu. The MOOD metrics set. In *Proceedings ECOOP Workshop on Metrics*, 1995.

[2] V. R. Basili. The role of experimentation in software engineering: past, current, and future. In *ICSE '96: Proceedings of the 18th international conference on Software engineering*, pages 442–449, Washington, DC, USA, 1996. IEEE Computer Society. ISBN 0-8186-7246-3.

[3] V. R. Basili, R. W. Selby, and D. H. Hutchens. Experimentation in software engineering. *IEEE Transactions on Software Engineering*, 12(7):733–743, 1986. ISSN 0098-5589.

[4] V. R. Basili, G. Caldiera, and H. D. Rombach. The goal question metric approach. In *Encyclopedia of Software Engineering*. Wiley, 1994.

[5] V. R. Basili, L. C. Briand, and W. L. Melo. A Validation of Object-Oriented Design Metrics as Quality Indicators. *IEEE Trans. Softw. Eng.*, 22(10):751–761, 1996.

[6] S. R. Chidamber and C. F. Kemerer. A Metrics Suite for Object-Oriented Design. *IEEE Transactions on Software Engineering*, 20(6):476–493, 1994.

[7] P. B. Crosby. *Quality is free : the art of making quality certain*. McGraw-Hill, New York :, 1979.

[8] G. D. and M. J. Why software quality matters: 'we did nothing wrong'. *Baseline*, Mar 2004.

[9] M. Ge and M. Helfert. A review of information quality research — develop a research agenda. In *Proceedings of the 12th International Conference on Information Quality*, November 2007.

[10] R. Harrison, S. J. Counsell, and R. V. Nithi. An Investigation into the Applicability and Validity of Object-Oriented Design Metrics. *Empirical Software Engineering*, 3(3):255–273, 1998.

[11] B. Henderson-Sellers. *Object-oriented metrics: measures of complexity*. Prentice-Hall, Inc., 1996.

[12] ISO. ISO/IEC 9126-1 "Software engineering - Product Quality - Part 1: Quality model", 2001.

[13] ISO. ISO 9000:2005 "Quality management systems - Fundamentals and vocabulary", 2005.

[14] J. Juran. *Juran's Quality Control Handbook*. McGraw-Hill, 5th edition, 1998.

[15] B. K. Kahn, D. M. Strong, and R. Y. Wang. Information quality benchmarks: product and service performance. *Commun. ACM*, 45(4):184–192, 2002.

[16] B. D. Klein, D. L. Goodhue, and G. B. Davis. Can humans detect errors in data? impact of base rates, incentives, and goals. *MIS Q.*, 21(2):169–194, 1997.

[17] S. T. Kresge. Aircraft accident investigation, f/a-22 s/n 00-4014. Technical report, 422nd Test and Evaluation Squadron Nellis Air Force Base (AFB), NEVADA, USA, Feb 2005.

[18] W. Li and S. Henry. Maintenance metrics for the object oriented paradigm. In *IEEE Proceedings of the First International Software Metrics Symposium*, pages 52–60, May 1993.

[19] R. Lincke, J. Lundberg, and W. Löwe. Comparing software metrics tools. In *Int. Symp. Software testing and analysis (ISSTA)*, pages 131–142, New York, NY, USA, 2008. ACM.

[20] R. Lincke, W. Löwe, and T. Gutzmann. Software Quality Prediction Models Compared. In *10th International Conference on Quality Software (QSIC 2010), Zhangjiajie, China*, July 2010.

[21] W. Löwe and T. Panas. Online-construction of software comprehension and visualization tools. *International Journal of Software Engineering and Knowledge Engineering – Special Issue on Maturing the Practice of Software Artifacts Comprehension*, 6(15):995–1025, 2005.

[22] J. A. McCall, P. G. Richards, and G. F. Walters. Factors in Software Quality. Technical Report Volume I, NTIS, NTIS Springfield, VA, 1977. NTIS AD/A-049 014.

[23] S. McConnell. *Code Complete, Second Edition*. Microsoft Press, Redmond, WA, USA, 2004.

[24] R. Y. Wang and D. M. Strong. Beyond accuracy: what data quality means to data consumers. *J. Manage. Inf. Syst.*, 12(4):5–33, 1996.

[25] A. Wingkvist, M. Ericsson, R. Lincke, and W. Löwe. A metrics-based approach to technical documentation quality. In *7th Int. Conf. Quality of Information and Communications Technology, Porto, Portugal*, Sep 2010.

[26] A. Wingkvist, M. Ericsson, W. Löwe, and R. Lincke. Incorporating information quality in software development. In *33rd Information Systems Research Seminar in Scandinavia, Rebild, Denmark*, Aug 2010.

[27] A. Wingkvist, W. Löwe, M. Ericsson, and R. Lincke. Analysis and visualization of information quality of technical documentation. In *4th European Conf. Information Mgmnt and Eval, Lisbon, Portugal*, Sep 2010.

[28] A. Wingkvist, M. Ericsson, and W. Löwe. A software infrastructure for information quality assessment. In *16th International Conference on Information Quality (ICIQ) , Adelaide, Australia*, Nov 2011.

[29] A. Wingkvist, M. Ericsson, and W. Löwe. Analysis and visualization of information quality of technical documentation. *The Electronic Journal of Information Systems Evaluation (EJISE)*, 14(1):150–159, 2011.

[30] A. Wingkvist, M. Ericsson, and W. Löwe. Making sense of technical information quality a software-based approach. *Journal of Software Technology: Analyzing and Measuring Information Quality*, 14(3):12–18, 2011.