



A validated model for the scoping process of quality requirements: a multi-case study

Thomas Olsson¹ · Krzysztof Wnuk² · Slinger Jansen^{3,4}

Accepted: 27 November 2020 / Published online: 2 March 2021
© The Author(s) 2021

Abstract

Quality requirements are vital to developing successful software products. However, there exist evidence that quality requirements are managed mostly in an “ad hoc” manner and down-prioritized. This may result in insecure, unstable, slow products, and unhappy customers. We have developed a conceptual model for the scoping process of quality requirements – QREME – and an assessment model – Q-REPM – for companies to benchmark when evaluating and improving their quality requirements practices. Our model balances an upfront forward-loop with a data-driven feedback-loop. Furthermore, it addresses both strategic and operational decisions. We have evaluated the model in a multi-case study at two companies in Sweden and three companies in The Netherlands. We assessed the scoping process practices for quality requirements and provided improvement recommendations for which practices to improve. The study confirms the existence of the constructs underlying QREME. The companies perform, in the median, 24% of the suggested actions in Q-REPM. None of the companies work data-driven with their quality requirements, even though four out of five companies could technically do so. Furthermore, on the strategic level, quality requirements practices are not systematically performed by any of the companies. The conceptual model and assessment model capture a relevant view of the quality requirements practices and offer relevant improvement proposals. However, we believe there is a need for coupling quality requirements practices to internal and external success factors to motive companies to change their ways of working. We also see improvement potential in the area of business intelligence for QREME in selecting data sources and relevant stakeholders.

Keywords Quality requirements · Non-functional requirements · Case study

Communicated by: Tao Yue

✉ Thomas Olsson
thomas.olsson@ri.se

Extended author information available on the last page of the article.

1 Introduction

Quality requirements (a.k.a. non-functional requirements) is one of the most researched areas within requirements engineering (Ambreen et al. 2018). Several studies conclude that quality requirements are essential, but not systematically handled (Berntsson Svensson et al. 2012; Ameller et al. 2016). Our previous work brings supporting evidence that addressing deficiencies in quality requirements take a long time (Olsson et al. 2019). The main reasons are a lack of explicit handling of quality requirements on a strategic level (also highlighted by Ameller et al. (2013), Eckhardt et al. 2016) and a lack of a feedback-loop in the scope decision process, which is one possibility to understand better the user experience and perception of the quality requirements. Companies working agile also struggle with managing quality requirements when looking at the big picture (team coordination challenge) and making unstated assumptions (conceptual challenge) (Alsaqaf et al. 2019).

This study continues our previous research efforts on understanding and supporting decision making about quality requirements. We have previously performed a longitudinal case study of one company's quality requirements decisions (Olsson et al. 2019). We empirically identified constructs underlying the scope decision process for quality requirements. Based on those constructs, we developed the QREME conceptual model (Olsson and Wnuk 2018). In this paper, we present Q-REPM, an operationalization of the QREME conceptual model for the scoping process of quality requirements. Furthermore, we validate Q-REPM and confirm QREME constructs' presence in a confirmatory multi-case study with five companies in Sweden and The Netherlands.

QREME is a conceptual framework (Olsson and Wnuk 2018) for decision making about quality requirements. The aim is to improve the integration of data-driven requirements engineering (Maalej et al. 2015) into the scoping process by clarifying which roles should be involved in the decision process. Scoping are the activities performed to identify which requirements should be part of a software release and the decision process around it (Wnuk and Kollu 2016). The idea with QREME is to bridge plan- and data-driven principles by utilizing competencies across roles in the companies. Q-REPM provides a benchmark instrument for decision-making about quality requirements that supports the introduction of more data-driven decision-making. Q-REPM is intended to help software companies understand the current practices and possible improvement areas.

We have conducted a multi-case study (Runeson et al. 2012) with five companies from Sweden and The Netherlands from different domains. The objective is to validate Q-REPM and the QREME constructs (Olsson and Wnuk 2018). We performed semi-structured interviews to understand how the companies work and workshops to validate our findings from the interviews and evaluate the usefulness of Q-REPM.

The paper is structured as follows: In Section 2, we present a summary of empirical studies on quality requirements. QREME and Q-REPM are presented in Section 3. We elaborate on our research questions and the case study protocol in Section 4. We interviewed 25 persons and held workshops with 30 participants at the five companies. The cases are presented in Section 5. Section 6 presents the results from the interviews and workshops. In Section 7, we discuss the results, lessons learned, and the improvement areas we see for QREME and Q-REPM. The main improvement area is the connection to business intelligence. We also discuss a rationale for using QREME, especially to broaden the input used for elicitation by using a data-driven approach – even if a company does not have direct users as their customers. The main threats to validity are construct validity, generalizability, and confirmation bias. Our analysis of validity threats and limitations are found in Section 8. Lastly, we conclude in Section 9 that our results with immature handling of quality requirements from the

case study are in line with other research. We conclude, however, Q-REPM can both reliably uncover this and provide guidance to improve the quality requirements practices.

2 Related Work

The presence of quality requirements varies, as reported in several empirical studies (Berntsson Svensson et al. 2013; Olsson et al. 2019; Calazans et al. 2019; Shahrokni and Feldt 2013). However, we see no clear trend nor clear analysis explaining the variation in different contexts. Furthermore, surveys of personal opinions arrive at similar results (Berntsson Svensson et al. 2012; Benslimane et al. 2007; Ameller et al. 2012; Daneva et al. 2014; De La Vara et al. 2011; Ameller et al. 2016; García-Mireles 2016; Caracciolo et al. 2014). An extensive interview study on quality requirements in the context of model-driven development finds that about half of companies prioritize quality requirements as important as functional ones and that 3 out of 4 of those come from the embedded systems domain (Ameller et al. 2019). The related work implies that presence and opinion on the priority of different quality requirements vary. This further implies that there is a need to steer and evaluate the scope of quality requirements continuously. Sveningson et al. conclude that there is a relationship between the road map, distance to the users, and experimentation (Sveningson et al. 2019), which is in line with our model (Olsson and Wnuk 2018). This paper presents a case study on how companies steer their scope process and cope with changes in priority.

Quality requirements have different sources. Architects are sometimes involved in the elicitation and definition of quality requirements (Ameller et al. 2012; Daneva et al. 2013). There are several papers on using different kinds of user reviews on mobile app markets as a potential source of quality requirements (Groen et al. 2017; Jha and Mahmoud 2019; Wang et al. 2018; Lu and Liang 2017). This is sometimes referred to as CrowdRE (Glinz 2019) or data-driven requirements engineering (Maalej et al. 2015). One study found that users are not sufficiently involved in the elicitation (Grimshaw and Draper 2001). We complement the existing work by collecting empirical data on the usage of different sources for quality requirements elicitation.

Overall, there is not much work on how quality requirements evolve over the product lifecycle. In our previous work, we studied the scope decision process over five years (Olsson et al. 2019). Ernst and Mylopoulos study 8 open source projects (Ernst and Mylopoulos 2010) on the fluctuations of priorities among quality requirements throughout the projects. The 8 projects had different trends in terms of how the priority changed over time, and the authors could not confirm their assumption of increasing importance of quality requirement over time. Ho et al. have published a study on the presence of “Not a Problem” issue reports compared to how precise quality requirements are written (Ho et al. 2008). The main result is that the more accurate the quality requirements, the minimize the “Not a Problem” issue reports. In this paper, we complement the existing work with empirical data on how information flows across roles and phases in the development at a point in time.

Software architects are often not involved in scoping of quality requirements, despite being – the primary source of quality aspects (Ameller et al. 2012; Daneva et al. 2013; Daneva et al. 2014). Our previous work found that relying on external stakeholders might lead to long lead-times and incomplete quality requirements (Olsson et al. 2019). We found one study on quality requirements handling in an Agile context. Alsaqaf et al. report, for example, that communication and unstated assumptions are significant challenges for quality requirements (Alsaqaf et al. 2019). Even though opinion surveys indicate that subjects claim to prioritize and explicitly work with quality requirements (Berntsson Svensson

et al. 2012; Benslimane et al. 2007; Ameller et al. 2012; Daneva et al. 2014; De La Vara et al. 2011; Ameller et al. 2016; García-Mireles 2016; Caracciolo et al. 2014), there are indications that implicit quality requirements handling are common, and this leads to misalignment. In this study, we study different roles at different companies to understand the alignment of the scope decision process.

3 Operationalizing QREME: Q-REPM

The terms quality requirements and non-functional requirements are the two prevailing terms. One can claim that many “non-functional” requirements are in fact functional (Eckhardt et al. 2016; Berntsson Svensson et al. 2013). As pointed out by Martin Glinz in 2007, there was no generally agreed definition of quality requirements (Glinz 2007). This paper uses the term quality requirements in line with Glinz’s definition: quality requirements are attributes or constraints.

QREME is a conceptual model for the scoping process of quality requirements (Olsson and Wnuk 2018). For QREME to be used, it needs to be operationalized. We, hence, created a framework called Q-REPM to assess the ways of working. Q-REPM is based on QREME and UNI-REPM (Svahnberg et al. 2015).

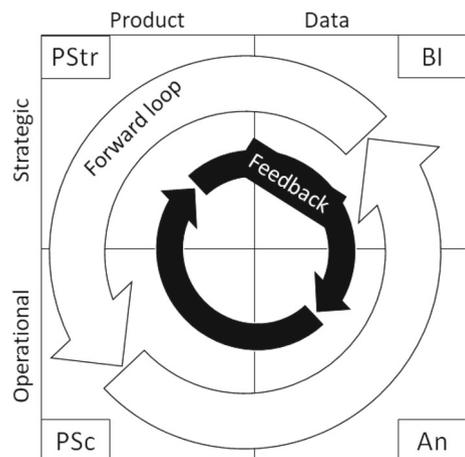
3.1 QREME

QREME is a conceptual model for the scoping process of quality requirements (Olsson and Wnuk 2018). QREME introduces the three constructs of the strategic and operational level, the product and data dimension, and the forward- and feedback-loop, see Fig. 1.

At the strategic level, scope decisions across multiple products – for a product portfolio or a product line – are made. Decisions are made for which quality requirements to include and high-level planning for when to use a forward-loop and when to use a feedback-loop. At the operational level, scope decisions for products and their releases are made. The decisions refine the strategic level into quality requirements and quality levels.

Scope decisions are separated from data, e.g., which competitor intelligence or usage data to collect. QREME introduces the product and data dimension to the scope decision

Fig. 1 The QREME Conceptual model (Olsson and Wnuk 2018). When combining the levels and dimensions, four scope decision areas are created (cf. Fig. 1): Product portfolio strategy (PStr), Product Scope (PSc), Business intelligence (BI) and Analytics (An)



process to more explicitly connect the two. Decisions on what to include in the product are made on the product dimension, whereas in the data dimension, decisions on what data to collect and analyze (Olsson and Wnuk 2018). Data decisions can also include decisions on what to experiment on, in a data-driven manner (Fagerholm et al. 2017).

The scope decision process is conventionally viewed as a top-down process from idea to inclusion and implementation (Regnell and Brinkkemper 2005). This is referred to as the forward-loop. QREME introduces an additional loop: the feedback-loop. Through usage data analysis (either through experimentation or without), identify both quality requirements and quality levels. Once candidates are identified, they are decided by a product manager role whether to include or not. All roles involved in the scope decision process take input from many sources and refine information and decisions. If a role, e.g. the product manager, sees a critical quality requirement to address which is not included in the strategy, there is a need for a process to handle this discrepancy. Both loops traverse the four scope decision areas in opposite directions and at different speeds (Olsson and Wnuk 2018).

When combing the levels and dimensions, four scope decision areas are created (cf. Fig. 1): Product portfolio strategy (PStr), Product Scope (PSc), Business intelligence (BI) and Analytics (An).

3.2 UNI-REPM

UNI-REPM is a requirements engineering process assessment and improvement framework (Svahnberg et al. 2015), bridging the gap between the theoretical world and practical reality. UNI-REPM is organized in areas and sub-areas, such as requirements elicitation and deliverable (Svahnberg et al. 2015). UNI-REPM defines actions, such as “identify and involve relevant stakeholders” or “Baseline quality levels”. The actions are assessed in the company’s requirements engineering practices through interviews and reviews. UNI-REPM, furthermore, defines a maturity level on actions: Level 1 Departure, Level 2 Intermediate, and Level 3 Destination. Before addressing actions on a higher level, all lower level actions should first be performed.

in UNI-REPM, each action is assessed as “Not performed”, “Performed” or “Satisfied/explained”.

- Not performed - actions that are not performed to a satisfactory level.
- Performed - actions that are considered to be performed to a satisfactory level.
- Satisfied / explained - actions that are not performed, but there is an acceptable explanation for it (not applicable).

When all actions on a level, e.g. Departure, are either performed or satisfied/explained, the collected rating of the organization is set to that level.

UNI-REPM is not always suitable and might also have deficiencies. This is evaluated through Satisfied/explained. Satisfied/explained implies the underlying theory does not apply to that company and can be seen as a deficiency or limitation in the model (Svahnberg et al. 2015). Hence, either the model is missing something or is not appropriate for the organization in question.

3.3 Q-REPM

Our implementation proposal uses QREME to enrich UNI-REP with additional process areas and actions for quality requirements decision making based on QREME. When designing Q-REPM, we went through the following steps:

1. We mapped the existing actions from UNI-REPM to QREME. For the actions relevant for QREME, we put them into PStr, PSc, BI, and An. For the actions relevant for either the levels – operational or strategic – they are put on the left side as applicable for that level. Similarly, for the dimensions – Product and Data.
2. Some of the existing actions in UNI-REPM are made redundant and replaced with new QREME derived actions, e.g., *RE.GA.a1 Elicit Quality Requirements*.
3. We identified actions in UNI-REPM that need to be changed, e.g., rename or made more explicit.
4. We define new actions to cover the remaining QREME aspects.

The assignment of an action to a particular level in Q-REPM is a constructive contribution. In Section 7.1, we elaborate a bit on the future work with respect to the level assignment.

The mapping can be found in Table 2 in the appendix. Changed actions are italic, and new actions are highlighted bold.

We applied a combination of inductive (open) (Pettersson et al. 2008) and prescriptive (model-based) assessment (Svahnberg et al. 2015) to explore how companies work with quality requirements. The assessment can be made through semi-structured interviews. The interview guide is based on Q-REPM. Open questions are asked for the different areas, covering the actions within the area.

Improvement recommendations are based on fulfilling all actions on the lower levels before addressing actions on a higher level. The exact practice to introduce to address a specific action is company-specific, based on the assessor's knowledge and experience.

4 Research Method

We have conducted a confirmatory multi-case study (Runeson et al. 2012) with 5 companies from Sweden and The Netherlands from different domains to validate Q-REPM and to confirm the presence of the QREME constructs (Olsson and Wnuk 2018).

Our study's *objective* is exploration of the usefulness of Q-REPM and confirmation of constructs (see Section 3) and propositions constituting QREME (Olsson et al. 2018, 2019). We study *scope decision process of quality requirements* via semi-structured interviews to understand how the companies work and workshops to validate our findings from the interviews and to evaluate the usefulness of Q-REPM. The companies were selected using convenience sampling with contrasting cases.

4.1 Research questions

Based on our research goal, we defined two research questions. The relationship between the research question and concepts are found in Fig. 2.

RQ1 How useful is the Q-REPM in practice?

RQ2 Are the underlying constructs of QREME observable in practice?

Regarding RQ1, we want to evaluate both if the framework is useful to assess decision making for quality requirements and that improvement recommendations based on Q-REPM are helpful for the companies. The former is achieved through the interviews and workshops as well as discussions among the researchers. The latter is performed by presenting recommendations to the companies and discussing with them in a workshop whether the

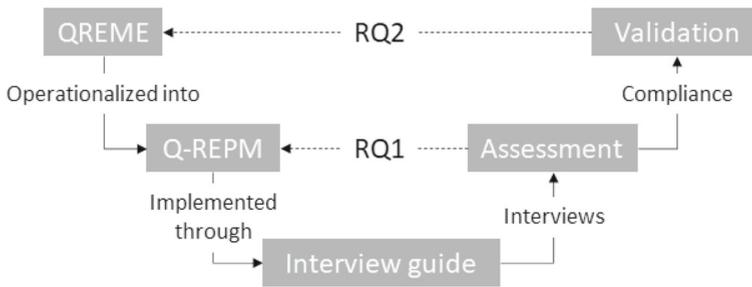


Fig. 2 The relationship between the QREME conceptual mode, Q-REPM, and the empirical study to answer the RQ1 and RQ2

recommendations are relevant in the companies' context; the improvements are actionable and realistic to introduce.

Based on the qualitative and quantitative data from RQ1, we trace back the results from five companies to the conceptual elements of QREME. This enables us to evaluate RQ2, whether we can find the same constructs at the case companies in this study, as in our previous work (Olsson et al. 2019), or whether QREME should be enriched with new constructs.

4.2 Selection strategy

We have focused on selecting a set of software-intensive product development companies that operate in different domains and have different years of experience within the software business. We focused on companies offering software products in an open market (MDRE) (Regnell and Brinkkemper 2005) because QREME constructs are grounded in the market-driven requirements engineering challenges and characteristics. We aimed to study companies with both short and long development cycles and various release frequency.

We focused on including strategic and operational level roles to surveying different perspectives (e.g., product manager and business intelligence) and several individuals from the same position to detect inconsistencies.

4.3 Process

We treated each of the five cases as individual case studies. We followed the same principle steps for each case.

1. **On-boarding meeting** – 1-hour meeting with a sponsor on the company side to get buy-in and to get them on-board. Purpose: Agreement on what to do. Outcome: contact person to help to find individuals and further planning and team to receive.
2. **Planning** – Work with the contact person to get basic information and to identify individuals to interview. Book meetings, etc.
3. **Tailor assessment material** – Tailor (if needed) terminology, the number of meetings, etc. Complement if specific topics are interesting for the companies.
4. **Interviews** – Elicit information on current practices.
5. **Wrap-up workshop** – Validate our understanding from the interviews and evaluate Q-REPM in terms of usefulness.
6. **Final report** – Finalize report to the company based on wrap-up session.

The steps 1-3 were not strict in the sense of timing and lead-time, as long as they happened before the interviews. The interviews were done in person, with one exception from FinComp where one interview was done over the phone. All the workshops were conducted as physical meetings.

4.4 Data Collection

We performed data collection through interviews and workshops for all five cases. The interviews were focused on eliciting information on how the companies work with quality requirements. The interviews were followed by a wrap-up workshop where the results from the interviews were presented, and additional feedback was elicited regarding the usefulness of Q-REPM and improvement suggestion proposed based on the assessment.

We performed interviews with either group of 2-4 individuals representing one role within their organization or individual interviews. The interview sessions lasted between 1-2 hours. The sessions consisted of open questions, as part of the inductive part of the assessment, as well as closed questions as part of the model-based assessment (see Section 3). We used the predefined questionnaire to guide the interviews. It consisted of open and closed interview questions. The interviews were structured as followed:

1. The first part focused on getting an understanding of the interviewees' perception of quality requirements (definition, which quality requirements are important, why, etc).
2. The second part focused on getting the context (product, maturity, process).
3. The third part focused on collecting the data related to Q-REPM (the process used to work with quality requirements). We used a combination of open and closed questions which cover the assessment sheet.

We gradually built up our understanding of the company's ways of working by discussing topics and specific details from previous interviews. In this way, several roles and interviewees could triangulate the results, improving reliability.

We summarized the interviews into a report which was sent to the companies before the wrap-up workshop. TransComp did not receive any material before the workshop as the interviews and workshop were held only one day apart. For the other cases, there were, for logistical reasons, 2-4 weeks between the interviews and workshop. The participants in the workshops were senior employees as appointed by the company and the interviewees (though not all participated, see Table 1). The structure of the workshop was are follows:

1. The validation was done as a presentation of the results from the interviews where the participants could comment on our findings.
2. The evaluation was done when the recommendations were presented. We elicited the opinions about the utility of the Q-REPM-based improvements. Lastly, the overall approach was discussed to evaluate whether the setup with interviews, workshops, etc., was sensible.

We sent each company a final report after the workshop, summarizing our findings.

4.5 Analysis

We performed a cross-case analysis (Seaman 1999). Qualitative data obtained during the interviews and workshops were classified according to Q-REPM. We classified each case individually. The classification was performed by one researcher and validated another participating researcher. We discussed disagreements and uncertainties until we reached

Table 1 An overview of the companies and interviewees. The scope was either a specific R&D unit, a business unit or the entire company. Some persons only participated in the interviews (Int.) or workshop only (WS)

ID	Size	Scope	#	Role	Domain
TranspComp	Large	R&D unit	5	Requirements engineer Scrum master Product manager Architect Line manager (WS)	Transportation
ISComp	Medium	Company	6	Developer Line managers (2) Product manager Architect (WS) Architect (Int.)	Information System
eCommComp	Small	Company	6	Project leader (Int.) Line managers Line managers (WS) Product manager Tech lead (2)	e-Commerce
RetailComp	Medium	Company	10	Line manager Line manager (Int.) Architect Developer (3 Int.) Consultants (2 Int.) Product manager (2)	Retail
FinComp	Large	Business unit	10	Line manager Line manager (3 WS) Architect (2) Developer Project leader (2) Product manager	Finance

a consensus. Furthermore, uncertainties in the classification of the interview notes were brought up in the workshop for discussion. We also recorded the contextual factors market domain, size of the company, applicable organizational scope. We complement the structured context factors with a general description of the companies to provide additional background.

The different cases were compared, and contextual factors were analyzed. Specifically, we wanted to understand if some actions in Q-REPM might only be relevant in a particular context, e.g., large companies. Furthermore, we also wanted to assess if there are actions that are not relevant. The context description is, therefore important to capture in sufficient detail so that this information later could be used to evaluate the applicability of the study's results in other contexts.

However, the recommendations are company-specific and as such, cannot be evaluated across the cases.

5 Case companies

37 persons from 5 different companies from Sweden and The Netherlands participated in our study. We ran 25 interviews with 31 interviewees and 5 workshops with 30 participants – 23 participants attended both an interview and a workshop, see Table 1.

TranspComp - TranspComp is one of the large companies and also the most traditional one in terms of how software is developed. TranspComp develops hardware and software for the transportation domain. Furthermore, they release new software seldom, and project lead-times are usually long. Much of the development is based on contracts with specific customers with a specification. Therefore, TranspComp has a strong emphasis on the forward-loop and almost completely lacking the feedback-loop. The implementation is done in smaller teams influenced by Scrum (Schwaber 2004); essentially, the lowest level of refinement is Agile but not the overall scope.

ISComp - ISComp develops a pure software product with a client-server architecture. The product is deployed on both physical dedicated servers and as a cloud solution. All but one of the interviewees mentioned ISO25010 (ISO 2011) at ISComp. Hence, the interviewees at ISComp have a good understanding of quality requirements. Despite this, they stated that *“Quality requirements are not prioritized.”* and *“We’re driven by customer requirements, so if they report a quality requirement we focus on that.”*

eCommComp - eCommComp is the smallest and youngest company in the study. Projects are in the range of weeks and months and not that many persons involved. Due to its small size, informal communication is utilized. Hence, more formalized practices and processes are often not necessary.

RetailComp - RetailComp develops an embedded software product in the retail domain. The hardware is mostly purchased, and RetailComp develops the software for the devices as well as the server solution for the system. Installation and customization at the customer sites are complex, and RetailComp provides a consultancy service for this.

FinComp - FinComp is a large company with a long history and global reach. However, the business unit in question is somewhat of a company within the company. They have also recently undergone a major reorganization and established a new way of working. They are establishing DevOps (Hüttermann 2012) practices, which are seen in the assessment results.

6 Analysis and results

6.1 RQ1 How useful is the Q-REPM in practice?

The companies perform only a small number of the suggested actions in Q-REPM (median 24%). This is seen in Fig. 3 by many black areas. The exception is the PSc actions – see Fig. 3. This indicates that quality requirements are handled well in the requirements scoping process at the operational level, e.g. in the backlog or documented requirements.

The empirical findings bring strong supporting evidence that very few of the actions in Q-REPM are irrelevant. Q-REPM contains 34 actions where 11 are part of UNI-REPM, one is changed from UNI-REPM and 22 actions are new; see Fig. 3. We assess that three

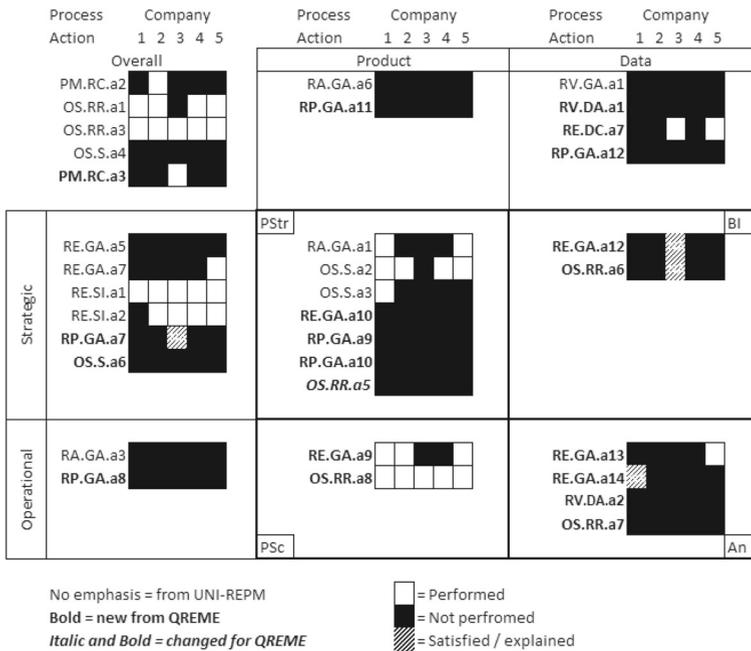


Fig. 3 Assessment for the companies. The new actions are in Bold and revised actions in italic. Details can be found in Table 2

actions are “Satisfied/explained” for eCommComp and one from TranspComp, see Fig. 3. This provides supporting evidence that Q-REPM is relevant to industry practice.

Q-REPM introduces scoping actions on a strategic level (PStr) and in the alignment between the operational and strategic levels, see Fig. 1. The companies perform 29% of the PStr actions and none of the general Product actions. This is in line with our observations from our previous study on quality requirements (Olsson et al. 2019). ISComp and RetailComp recognized deficiencies in handling quality requirements that confirm our assessment. We believe that the lack of strategic guidance in quality requirements is a root-cause of many deficiencies. However, without a longitudinal study where we can observe the changes from improving the strategic scoping actions leaves us with logical assumptions of the perceived benefits. Hence, we need more empirical studies to understand this in more detail.

Likewise, we observe that none of the companies perform the data dimension – Data, BI, and An – to a satisfactory level. eCommComp and FinComp perform *RE.DC.a7 Elicit information about System’s measurement capability* and FinComp also performs *RE.GA.a13 Elicit quality requirements on data usage level*. All companies except TranspComp use log files and some analytics, albeit not systematically. Hence, we assess them as not performing the actions related to data collection at an operational level. FinComp – who is actively trying to establish DevOps – systematically collects data for quality requirements - the only in our study. None of the companies perform actions with data on the strategic level. Working data-driven and collecting data from product usage are practices that have gained much interest in the last couple of years (Maalej et al. 2015). Hence, it is surprising that our companies are not more systematic in their measurement work. In our previous work, we observed that the scoping process evolution went from a conventional process with experts

as input for requirements through establishing a measurement program to an updated strategic process – albeit still forward-focused (Olsson et al. 2019). It seems as if TranspComp is still to start this journey, whereas ISComp and RetailComp have started it, and FinComp is the furthers along. We hypothesize that the combination of trying to establish DevOps practices and a revenue-sharing model are the two contributing factors why FinComp is more systematic in working data-driven. eCommComp does not have a systematic data collection program. However, given the small size, we assess that the strategic discussion and processes for the data dimension do not apply to them. Therefore, the assessment is marked as “Satisfied / explained”.

None of the companies systematically perform the BI actions, see Fig. 3. However, all companies, except eCommComp, do have explicit business intelligence activities, but not directly seen in the scoping work of quality requirements. Instead, the information from the marketing organization – typically responsible for business intelligence – is communicated to product managers and line managers. Hence, BI is only indirectly influencing requirements scoping. That is, there is an ad-hoc alignment, but it is not systematic nor explicit. eCommComp, being a small company and working mostly project-based, does not have explicit BI activities. Therefore, the actions for eCommComp are considered not applicable. This, again, confirms our previous results where BI was not an integral part of the scoping process (Olsson and Wnuk 2018).

None of the companies is using A/B testing to evolve products (RE.GA.a14). The companies share the common concern of alienating the users or that their customers are not the users (Rissanen and Münch 2015; Sveningsson et al. 2019). The fact that none of the companies in the study sell their products directly to the end-users seems to be a mental obstacle to engage in direct end-user interaction. One interviewee from FinComp commented that since Company FinComp does not know the users’ business process, it is not sensible to utilize techniques such as A/B testing. This is similar to the results of Rissanen and Münch (2015). We believe that this is a fallacy where the companies rely too much on their customers who might have other interests than the needs of the end-users. The companies become distant from the actual users and will have problems understanding the actual customer, e.g. through data-driven techniques.

6.2 RQ2 Are the underlying constructs of QREME observable in practice?

The core constructs of QREME are the strategic and operational levels, the product and data dimensions, and the forward- and feedback-loops – see Section 3. We argue that we have identified the constructs in all the companies, albeit instantiated quite differently.

6.2.1 Strategic and operational level

TranspComp – where the ways of working are forward-focused – invests significant effort on upfront analysis. There is also explicit work on a strategic level through a portfolio plan. In the portfolio plan, high-level concepts are documented, which serves as input in the customer dialogue. On an operational level, there is a combination of a conventional requirements specification as well as Scrum backlogs. We reason that the strategic level might be even more important in cases like TranspComp. As the lead-time from delivery to customer and experience from usage is longer, it is imperative to get it right from the start. Therefore, we conclude that a clear separation between the strategic and operational levels and structured dialogue with stakeholders on the strategic level is the necessary enablers for improved handling of quality requirements.

ISComp and RetailComp have a product base for which they adapt and deliver customer-specific variants. They do have explicit and systematic road map work on a strategic level. At the same time, both companies perceive that they are highly customer-driven and adapt the scope of releases according to customer wishes. Similarly, FinComp has a road map and, at the same time, works with vital customers who tend to be prioritized before the road map. The difference, however, is that the development is less contract-driven in terms of scope. Rather, the scope is mostly decided from new requirements submitted continuously rather than in contract-driven releases.

6.2.2 Product and data dimension

TranspComp, as mentioned, has projects with long lead-time and infrequent software updates. Hence, it is difficult to measure product usage and have that impact the scope of the product. This implies that the product and data dimension construct will be different for companies in a similar situation. We speculate that for TranspComp, systematically collecting product usage data is still relevant. However, instead of feeding it back to the same product that is being measured, the data can be used for future products.

RetailComp is starting a measurement program focused on key performance indicators at software update. However, it is still quite immature and under roll-out. An interviewee commented that *“If we start with analytics, we need to have a strategy”*. Hence, they are aware and thinking about working more data-driven, but is not yet moving in that direction. FinComp actively and systematically collects product usage data and is using it to define quality requirements that are funneled back through the scope flow. For example, they have instrumented their products to measure response-time in various use cases. Hence, FinComp is utilizing a basic feedback-loop. One interviewee commented that *“Some inconsistencies in quality requirements from different stakeholders”*. Essentially, different customers – which is not the same as users – have a different idea of which quality requirements are important and what the requested levels of quality are.

There are differences in how easy it is to collect and use product usage data for different products and services. A cyber-physical system where either capacity of the devices or the network capabilities are limited will make it more difficult to collect usage data than an e-commerce web site where hardware and network are rarely a limiting factor. TranspComp’s products are not connected to the internet, which means any usage or log data can only be retrieved when physically plugged into the products.

6.2.3 Forward- and feedback-loop

All of the companies have a forward-loop. This is judged by the presence of product managers (though the name varies) and their responsibility to define a road map and steer the direction of the implementation. The road map is refined in different ways – either by the product manager or other roles e.g., Scrum product owners. Any missed user expectations are funneled back to development through the issue flow and up the refinement chain if needed. Hence, it is reactive feedback from the forward-loop.

We observe that none of the companies perform “RE.GA.a5 Use Appropriate Elicitation Techniques according to Situation” and only FinComp is systematically performing “RE.GA.a7 Create Elicitation Channels for Requirements Sources”. For the QREME extension, this is not a static decision. Rather, this decision should be made continuously and especially whether to use the forward- or the feedback-loop.

TranspComp – which has long projects and years between software updates – is almost entirely forward-focused. This is not necessarily a problem in their current business envi-

ronment. We speculate that the feedback-loop defined today in QREME does not apply to TranspComp.

FinComp is the only company that systematically collects usage data. The focus is much on the technical performance of their product rather than understanding the users. The other companies in the study do collect various logs and usage data, but not systematically.

An interviewee from RetailComp comments that “*Maintenance is a big issue for us from a quality requirement perspective*”. Essentially, as the scope of the software grows, the quality levels e.g., performance decrease. Even though this is subjectively seen, it tends to degrade until it reaches a critical point when customers are no longer accepting a software release. RetailComp has introduced a measurement program to visualize essential quality requirements.

7 Discussion

Q-REPM was developed to support the integration of data-driven requirements engineering into decision making about quality requirements. Looking at the main goals of Q-REPM, we believe that it delivers an efficient benchmark instrument and identifies many improvement areas. When we consider that the companies only perform, in median, 24% (see Fig. 3) of the actions in Q-REPM, the companies seem immature in their quality requirements practices.

Quality requirements should be part of the product strategy scope (Kittlaus and Fricker 2017). However, the companies involved in our study are not performing strategic level actions. We speculate that this can be caused by a combination of the product lifecycle and market maturity. We have seen that early in the lifecycle, there will be less focus on quality requirements (Olsson et al. 2019). If the market is immature, we believe there are fewer but stronger customers who want their specific requirements fulfilled and who expect the quality to be implicit. However, if either the product lifecycle is further progressed or the market maturity is higher, we hypothesize that more attention will be paid to the quality requirements. Lastly, in a mature market with a mature product, the quality will play a differentiating role. We believe QREME can be a tool to initiate strategy changes. Suppose the analysis of the product lifecycle and the market analysis is combined with an assessment of Q-REPM. In that case, QREME can point to a wanted position in terms of quality requirements engineering (Table 2).

7.1 Improving Q-REPM

We find the mapping of QREME onto UNI-REPM to be straight-forward. However, there are several ways how QREME can be mapped onto UNI-REPM. We approach this by iteratively switching between a constructive design phase and an evaluation phase where Q-REPM is evaluated. We believe it is not possible to analytically design Q-REPM without empirically testing it and learning from the experience. We want QREME and Q-REPM to be useful, even though we will never be able to prove them *correct*¹.

Due to the large number of “Not performed” actions, we could not assess whether our assignment of actions into the three levels “Departure”, “Intermediate”, and “Destination” reflected what the companies and we (as assessors) judged to be the the most relevant

¹As George Box put it: “All models are wrong but some are useful”.

actions to address first. We see a need for more empirical work to gather more data to be able to assess this aspect of Q-REPM.

The forward- and feedback-loops are constructs in QREME. We have not yet mapped the actions to the loops. We have not yet mapped the actions to the loops. We speculate that this might both reveal additional actions as well as lead to revision of some of the actions. We plan this for the next iteration of QREME.

It was no surprise that the data part of QREME required more new actions than the other parts, as data-driven requirements engineering is an emerging topic. However, when we reflect on Q-REPM, we see that the PSc part only has two actions, both new, see Fig. 3. Our perception is that companies invest most in the PSc of quality requirements engineering. Also, we argue that most requirements engineering literature also focuses on the PSc part. Lastly, we believe that many Agile methods are focused on this part. We conclude that the PSc part of Q-REPM needs further elaboration to be able to sufficiently assess a company's ways of working. We plan to study Agile methods in more detail for which actions they propose as part of PSc in QREME to see if QREME and Q-REPM could be improved based on them. Furthermore, we see a need to empirically explore how companies handle quality requirements in more detail in the PSc part.

When we performed the interviews and workshops, the BI part of Q-REPM was the most difficult for us to explain. We believe the reasons are: 1) it is the most unusual part to include in the scoping process, 2) it is the most immature part of Q-REPM as it is the furthest away from our research areas as well. We believe that with data-driven requirements engineering emerging as a new area and the need to understand customers and users on a much broader scale than simple sampling, through e.g., focus groups and expert opinion, it will increase the importance of BI in Q-REPM. We see the need for future research in this area, from a software engineering perspective as well as business and marketing.

Some actions in Q-REPM are easily qualified – e.g., “OS.S.as Define Product Strategies” – whereas others are more of a judgment – e.g., “RE.GA.a13 Elicit quality requirements on a data level”. For the former, we can simply check the presence of the artifact. For the latter, it is more of a qualitative judgment from the interviews. We do not think it is reasonable to make all actions easier to be judged. However, we believe that if we add example artifacts and tools to Q-REPM, we would be able to improve the assessment.

7.2 Lessons learned

The main lesson that we learned from the low performance of the studied companies, see Fig. 3, is that managing quality requirements remains challenging and data-driven approaches only amplify this challenge. eCommComp and RetailComp systematically perform 6 of the actions, TranspComp and ISComp 8 actions, and FinComp 11 actions. We believe Q-REPM has captured a representative picture of the companies' ways of working with quality requirements. Hence, Q-REPM seems able to identify deficiencies according to the model. In that sense, Q-REPM works. The companies also confirmed that the improvement suggestions, based on the assessment and our experience, are relevant in their context and realistic to implement in their ways of working.

There are different aspects of quality requirements, such as security and usability. The idea with QREME is to incorporate a situational aspect in how specific requirements are handled – through a forward-loop or feedback-loop. However, when discussing with the companies, we often discussed that different quality requirements necessitate different handling. At the same time, we do not believe in large and complicated processes with numerous alternatives. Rather, we believe in using the knowledge and experience of different roles and supporting them in their alignment. At the same time, there might be recurring patterns

– e.g., security benefits more from a forward-loop than a feedback-loop and the other way around for usability – which are applicable in many contexts. FinComp, for example, had a separate group for security-related requirements. This group is part of the scoping process and both contribute with a review of requirements as well as proposes new requirements to include in the scope. This warrants further research.

In the final workshop, eCommComp commented that “*QREME is too elaborate for us today, but as we want to grow, we want to change our way of working to prepare for the future*”. eCommComp is the smallest in the study. TranspComp, on the other hand, is different than the other companies in that they release infrequently and has long lead-time projects. Q-REPM was more difficult to use for these two cases. For the former, a clearer guide of which steps to take first can help. For the latter, it might simply be that the data part of Q-REPM does not apply to the way it is defined right now. We believe the core of QREME should be general. However, QREME might not always be useful and the levels in Q-REPM might depend on the context. Regarding the former, in a waterfall-like development environment, for example in a safety-critical domain, working data-driven might not be possible. Hence, even if QREME can be used, it might not be useful. For the latter, the order in which actions are addressed – guided by the levels – might be context-dependent. We speculate that an adaptation of the levels could be connected to success factors from similar companies. We plan to explore the relationship between quality requirements and product success in future work.

7.3 Motivating QREME

A participant in the final workshop at FinComp commented that “*When working B2B, it is more difficult to work data-driven as we do not understand the complete use case, only our part*”. By “B2B”, the participant wanted to express that the users are not the customers and that the customers are adding other systems to the end-result the users get. Hence, FinComp only sees a part of the use case. It is true for many companies unless there is complete control of the value chain from end to end. It is also related to how independent the products or services are. For example, if the product is a car stereo, your main customers will be the car manufacturers – B2B. However, the stereo is quite self-contained in the car. Hence, we argue that it is still very relevant for a car stereo developer to work actively with data-driven approaches. On the other hand, engine control developers deliver a component that might be difficult to understand in isolation. We speculate that it might be a fallacy to argue that it is more difficult to work data-driven if you do not understand the complete use case of the users. Rather, we believe this as a defensive reaction to changes that might threaten the authority of the participant should they work more data-driven. This is similar to the “assumption trap”, where actors in an ecosystem assume to understand other actors added value (Bosch and Holmström Olsson 2018).

An interviewee from ISComp commented that “*We don’t care about existing customers*” when discussing the forward- and feedback-loop. Our experience of working with other companies is that this perception is quite common. That is, it is easy to focus on getting new customers and thereby focusing the scoping efforts on those stakeholders (Olsson et al. 2019). Hence, we conclude that there is a lack of interest in understanding the existing users and customers overall. We believe this is a risky and wasteful approach. It is risky because ISComp risks losing customers. It is wasteful as the product usage is fairly inexpensive to collect and yet has a wealth of information on the users – often a diverse group that is difficult to holistically understand with upfront sampling methods such as workshops, focus groups, market analysis, etc. We believe the ability to utilize product usage data and

other analytics is highly related to the overall maturity of a company. We are confident that Q-REPM can pinpoint these immaturities in this area of quality requirements.

We see a risk that if relying too much on stakeholders – who might have their requirements – there is a risk that the users' requirements are missed. Furthermore, the stakeholder disagreement illustrates another important aspect – the quality level will be different for different contexts and quality requirements are prioritized differently by different users and stakeholders. In our view, the implication is that working data-driven – by systematically collecting and analyzing usage data and by experimentation such as A/B testing – is crucial to get a realistic view of what the users perceive and need.

8 Threats to validity and limitations

Threats to construct validity (Runeson et al. 2012) are mostly touching upon the operationalization of QREME into Q-REPM. We have made mapping from the conceptual model with the actions and assessment questions. There are possible other conceptualizations of QREME into assessment questions or measures. For example, one could use only quantitative measures to investigate an example company. We used interviews during the assessment. This leaves us with a risk of misinterpretation (Runeson et al. 2012). We mitigated this by always having two researchers present in the interviews and by asking clarifying questions. The interview findings were further discussed in a workshop with the companies.

Study participation was voluntary and supported by a general interest in understanding quality requirements. Therefore, selection bias remains a validity threat of this study and negatively impacts the external validity of the findings (Runeson et al. 2012). We, however, argue the companies are representative of software development companies in their domains (Seddon and Scheepers 2012). Furthermore, as we had planning meetings with a company liaison to get suitable interviewees, we believe we have a representative sample of people and projects (Seaman 1999). All of the companies in our case study are influenced by Agile thinking (short iterations, close collaboration between product management and developers, self-managing teams) in one way or another - albeit the development cycle at TranspComp are long and not a typical Agile context. This is in line with our experience with other companies. All companies except eCommComp have for many years developed software based on a code-base that has evolved over many years. The software is a major part of the engineering efforts of the products. Furthermore, as pointed out by (Flyvbjerg 2006), the threats to generalization should not be exaggerated. As we have covered different sizes, several domains, and companies both in Sweden and The Netherlands, we believe the findings to be, at least to some extent, relevant for software-intense companies.

The sessions consisted of open questions, as part of the inductive part of the assessment, as well as closed questions as part of the model-based assessment (see Section 3). The flexible length of the interviews may pose some validity concerns. On one hand, shorter interviews could not provide the necessary depth into the discussed concepts and longer interviews may cause unnecessary fatigue and saturation among the interviewees. All researchers involved in this work have experience in performing interviews. This experience played a key role in deciding the suitable length of the interview. We did not experience any of the two before-mentioned issues during the interviews. However, we adjusted the interview length according to the needs and goals of the research.

As the researchers performing the evaluation also participated in the development of QREME, there is a risk of confirmation bias making the results less reliable. We mitigated this by always having two researchers present in the interviews, by having predefined ways to interpret the interviews through the model-based assessment part of Q-REPM, and by combining interviews and workshops to capture as unbiased input from the companies as

possible (see Section 4.5). Furthermore, the different cases were always performed by two researchers, including the analysis (the four-eyes principle). Hence, we addressed the threats to reliability (Runeson et al. 2012) by having multiple researchers part of all steps and by confirming our understanding of the workshops with the companies. Furthermore, when we experienced inconsistent views on whether an action was performed or not, we looked at additional sources, e.g., process documents or specifications, for additional evidence.

9 Conclusion

Anecdotally, deficiencies in quality requirements implementation are causing loss of business for companies. We see this in our results as well; ISComp and RetailComp see quality requirements leading to problems in timely deliveries or operations. However, TranspComp and FinCom – and to some extent eCommComp – seem to be performing well despite clear immaturity in the way they work with quality requirements. We cannot from this study conclude that quality requirements practices need to improve and that a new solution called QREME is needed. However, what we see in literature and our work, is a perception from many companies that quality requirements are not handled appropriately, yet neither they nor we can put a value on it.

We conclude that QREME was relevant and useful. The conceptual model guides how companies can work with quality requirements as well as can act as an assessment instrument – through Q-REPM – where practices can be benchmarked. Performing the assessment does not cost weeks of work and yet we identified several deficiencies in the companies' ways of handling quality requirements. Furthermore, the companies confirmed our interpretations and found the recommendations useful. Especially, we believe there is a need to broaden the scope of quality requirements to data-driven approaches. Our results indicate that the constructs of QREME can be used to address that particular issue. However, QREME needs to be validated in more cases to further improve the external validity.

We see two main streams of research going forward. Firstly, we would like to connect success factors to QREME and Q-REPM. The challenge is to isolate the influence of quality requirements specifically on success factors such as sales, customer satisfaction, or profit. There might also be internal success factors. For example, employee satisfaction with the process, internal alignment, and ability to deliver on time. However, despite many years of research on quality requirements, the interest seems to be fading. One reason might be that it is not clear if proper handling of quality requirements lead to tangible results. Hence, we want to explore other possible “architectural explanations” (Wieringa and Daneva 2015) and explore if there are conceptual constructs connecting quality requirements and success factors.

Secondly, several parts of QREME need further work. There are several ways QREME can be operationalized. Our approach is to iterate a constructive design phase with an empirical evaluation to refine and improve Q-REPM. Connecting the constructs and associated actions and practices from requirements engineering to marketing is an important area. We plan to explore the relationship between the data dimension of QREME and the selection of stakeholders and data-sources. We believe that accurately selecting the requirements sources has implications on the quality needs of a product and thus translates into the product success. Furthermore, we also plan to study and clarify different roles and their relationship to the QREME constructs. Another is to broaden the ideas for additional context e.g., longer lead-time projects. Lastly, we also need to further evaluate the data dimension of QREME. The companies only perform, in median, 24% of the actions in QREME; we need to empirically evaluate whether the actions and constructs from QREME are applicable in practice in companies more mature in their quality requirements handling.

Table 2 Q-REPM 1.2 Process areas and actions – see Section 3

ID	Q-REPM Action	Level	QREME	Company Transp Comp	IS Comp	eComm Comp	Retail Comp	Fin Comp
RE	Requirements Elicitation	H						
RE.GA	General Actions	H						
RE.GA.a5	Use Appropriate Elicitation Techniques acc. to Situation	2	Strat	N	N	N	N	N
RE.GA.a7	Create Elicitation Channels for Requirements Sources	3	Strat	N	N	N	N	Y
RE.GA.a9	Elicit Quality Requirements on operational scope level	1	PSc	Y	Y	N	N	Y
RE.GA.a10	Elicit Quality Requirements on product strategy level	1	PStr	N	N	N	N	N
RE.GA.a12	Elicit quality requirement on business intelligence level	2	BI	N	N	NA	N	N
RE.GA.a13	Elicit quality requirements on data usage level	2	An	N	N	N	N	Y
RE.GA.a14	Elicit quality requirements to experiment on	2	An	NA	N	N	N	N

Table 2 (continued)

ID	Q-REPM Action	Level	QREME	Company Transp Comp	IS Comp	eComm Comp	Retail Comp	Fin Comp
RE.SI	Stakeholder and Requirements Source Identification	H						
RE.SI.a1	Identify and Involve Relevant Stakeholders	I	Strat	Y	Y	Y	Y	Y
RE.SI.a2	Identify Other Requirements Sources	I	Strat	N	Y	Y	Y	Y
RE.DC	Domain Consideration and Knowledge	H						
RE.DC.a7	Elicit information about System's measurement capability	I	Data	N	N	Y	N	Y
DS	Documentation and Requirements Specification	H						
DS.GA	General Actions	H						
DS.DD	Documentation Deliverables	H						
RA	Requirements Analysis	H						

Table 2 (continued)

ID	Q-REPM Action	Level	QREME	Company Transp Comp	IS Comp	eComm Comp	Retail Comp	Fin Comp
RA.GA	General Actions	H						
RA.GA.a1	Perform Requirements Risk Analysis	1	PStr	Y	N	N	N	Y
RA.GA.a3	Analyze Requirements Relations	2	Oper	N	N	N	N	N
RA.GA.a6	Perform Refinement and Abstraction of Requirements	2	Product	N	N	N	N	N
RA.QA	Quality attributes analysis	H						
RA.PS	Problems and solutions analysis	H						
RP	Release Planning	H						
RPGA	General Actions	H						
RPGA.a7	Align product strategy with business intelligence	1	Strat	N	N	NA	N	N

Table 2 (continued)

ID	Q-REPM Action	Level	QREME	Company Transp Comp	IS Comp	eComm Comp	Retail Comp	Fin Comp
RPGA.a8	Align operational product scope with data usage	2	Oper	N	N	N	N	N
RPGA.a9	Create Basic QR road map and release plan	2	PSIr	N	N	N	N	N
RPGA.a10	Synchronize QR release plan with product road map	3	PSIr	N	N	N	N	N
RPGA.a11	Align product strategy with operational product scope	1	Product	N	N	N	N	N
RPGA.a12	Align business intelligence with data usage	2	Data	N	N	N	N	N
RPS	Requirements Selection	H						
RV	Requirements Validation	H						
RVGA	General Actions	H						
RVGA.a1	Validate requirements with relevant stakeholders	1	Data	N	N	N	N	N

Table 2 (continued)

ID	Q-REPM Action	Level	QREME	Company Transp Comp	IS Comp	eComm Comp	Retail Comp	Fin Comp
RVDA	Data Actions	H						
RVDA.a1	Baseline quality levels	1	Data	N	N	N	N	N
RVDA.a2	Establish a measurement program for relevant QRs	1	An	N	N	N	N	N
OS	Organizational Support	H						
OS.GA	General Actions	H						
OS.RR	Roles and Responsibilities	H						
OS.RR.a1	Assign Owner(s) of Requirements Dev. and Mgmt Processes	2	QREME	Y	Y	N	Y	Y
OS.RR.a3	Define Roles and Responsibilities for Release Planning	2	QREME	Y	Y	Y	Y	Y
OS.RR.a5	Define Roles and Responsibilities for Product Strat level	2	PStr	N	N	N	N	N

Table 2 (continued)

ID	Q-REPM Action	Level	QREME	Company Transp Comp	IS Comp	eComm Comp	Retail Comp	Fin Comp
OS.RR.a6	Define Roles and Responsibilities for Data Strat level	2	BI	N	N	NA	N	N
OS.RR.a7	Define Roles and Responsibilities for Data operational level	2	An	N	N	N	N	N
OS.RR.a8	Define Roles and Responsibilities for Product operational level	2	PS ^c	Y	Y	Y	Y	Y
OS.S	Strategies	H						
OS.S.a2	Define Product Strategies	2	PS ^{tr}	Y	Y	N	Y	Y
OS.S.a3	Define Product road maps	2	PS ^{tr}	Y	N	N	N	N
OS.S.a4	Communicate Strategies in Organization	3	QREME	N	N	N	N	N
OS.S.a6	Define strategies for selecting forward and feedback loop	2	Strat	N	N	N	N	N
PM	Requirements Process Management	H						

Table 2 (continued)

ID	Q-REPM Action	Level	QREME	Company Transp Comp	IS Comp	eComm Comp	Retail Comp	Fin Comp
PM.GA	General Actions	H						
PM.RM	Requirements Management	H						
PM.RM.a5	Establish a measurement program for relevant QRs	1	An	N	N	N	N	N
PM.RT	Requirements Traceability Policies	H						
PM.RC	Requirements Communication and Negotiation	H						
PM.RC.a2	Obtain common understanding of requirements among different involving roles	2	QREME	N	Y	N	N	N
PM.RC.a3	Establish Effective Communication With Decision roles	1	QREME	N	N	Y	N	N

Bold Id's are new actions in Q-REPM not present in UNI-REPM. Bold and italic Id's are changed actions. Level refers to process maturity level in Q-REPM (1 - Basic, 2 - Intermediate, 3 - Advanced) (see also (Svahnberg et al. 2015)). H in level indicates process area (as some areas of UNI-REPM are not updated, some process areas are empty in Q-REPM). If an action is performed at a company, it is marked as "Y"; if not performed "N", and if satisfied/explained as "N/A"

Acknowledgements We thank the companies for participating in the study and the valuable feedback we received. We would like to acknowledge that this work was supported by the KKS foundation through the S.E.R.T. Research Profile project at Blekinge Institute of Technology.

Funding Open Access funding provided by RISE Research Institutes of Sweden

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alsaqaf W, Daneva M, Wieringa R (2019) Quality requirements challenges in the context of large-scale distributed agile: An empirical study. *Inf Softw Technol* 110:39–55
- Ambreen T, Ikram N, Usman M, Niazi M (2018) Empirical research in requirements engineering: trends and opportunities. *Requir Eng* 23(1):63–95
- Ameller D, Ayala C, Cabot J, Franch X (2012) How do software architects consider non-functional requirements: An exploratory study. In: 2012 20th IEEE International Requirements Engineering Conference (RE). IEEE, pp 41–50
- Ameller D, Ayala C, Cabot J, Franch X (2013) Non-functional requirements in architectural decision making. *IEEE Softw* 30:61–67
- Ameller D, Galster M, Avgeriou P, Franch X (2016) A survey on quality attributes in service-based systems. *Softw Qual J* 24(2):271–299
- Ameller D, Franch X, Gómez C, Martínez-Fernández S, Araujo J, Biffi S, Cabot J, Cortellessa V, Méndez-Fernández D, Moreira A, Muccini H, Vallecillo A, Wimmer M, Amaral V, Böhm W, Bruneliere H, Burguño L, Goulão M, Teufl S, Berardinelli L (2019) Dealing with non-functional requirements in model-driven development: A survey. *IEEE Transactions on Software Engineering*
- Benslimane Y, Cysneiros LM, Bahli B (2007) Assessing critical functional and non-functional requirements for web-based procurement systems: a comprehensive survey. *Requir Eng* 12(3):191–198
- Berntsson Svensson R, Gorschek T, Regnell B, Torkar R, Shahrokni A, Feldt R (2012) Quality requirements in industrial practice: an extended interview study at eleven companies. *IEEE Trans Softw Eng* 38(4):923–935
- Berntsson Svensson R, Olsson T, Regnell B (2013) An investigation of how quality requirements are specified in industrial practice. *Inf Softw Technol* 55(7):1224–1236
- Bosch J, Holmström Olsson H (2018) Ecosystem traps and where to find them. *J Softw Evol Process* 30(11):e1961. <https://doi.org/10.1002/smr.1961>
- Calazans TSA, Paldés RA, Canedo ED, Masson TSE, Guimares FA, Rezende MFK, Braosi E, Kosloski RAD (2019) Quality requirements: Analysis of utilization in the systems of a financial institution. In: 2019 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM). IEEE, pp 1–6
- Caracciolo A, Lungu MF, Nierstrasz O (2014) How do software architects specify and validate quality requirements? In: European Conference on Software Architecture. Springer, pp 374–389
- Daneva M, Buglione L, Herrmann A (2013) Software architects- experiences of quality requirements: what we know and what we do not know? In: International Working Conference on Requirements Engineering: Foundation for Software Quality. Springer, pp 1–17
- Daneva M, Marczak S, Herrmann A (2014) Engineering of quality requirements as perceived by near-shore development centers' architects in eastern europe: the hole in the whole. In: Proceedings of the 8th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement. ACM, pp 19
- De La Vara JL, Wnuk K, Berntsson-Svensson R, Sánchez J, Regnell B (2011) An empirical study on the importance of quality requirements in industry. In: SEKE, pp 438–443

- Eckhardt J, Vogelsang A, Fernández DM (2016) Are non-functional requirements really non-functional? an investigation of non-functional requirements in practice. In: International Conference on Software Engineering. IEEE, pp 832–842
- Ernst NA, Mylopoulos J (2010) On the perception of software quality requirements during the project lifecycle. In: Requirements Engineering: Foundation for Software Quality. Springer, pp 143–157
- Fagerholm F, Sanchez Guinea A, Mäenpää H, Münch J (2017) The RIGHT model for Continuous Experimentation. *J Syst Softw* 123:292–305
- Flyvbjerg B (2006) Five misunderstandings about case-study research. *Qual Inq* 12:219–45
- García-Mireles GA (2016) Identifying relevant product quality characteristics in the context of very small organizations. *Comput Sci Inf Syst* 13(3):875–900
- Glinz M (2007) On non-functional requirements. In: 15th IEEE International Requirements Engineering Conference (RE 2007). IEEE, pp 21–26
- Glinz M (2019) Crowdre: Achievements, opportunities and pitfalls. In: 2019 IEEE 27th International Requirements Engineering Conference Workshops (REW), pp 172–173
- Grimshaw DJ, Draper GW (2001) Non-functional requirements analysis: deficiencies in structured methods. *Inf Softw Technol* 43(11):629–634
- Groen EC, Koczyńska S, Hauer MP, Krafft TD, Doerr J (2017) Users—the hidden software product quality experts?: A study on how app users report quality aspects in online reviews. In: 2017 IEEE 25th International Requirements Engineering Conference (RE). IEEE, pp 80–89
- Ho C-W, Williams L, Robinson B (2008) Examining the relationships between performance requirements and ‘not a problem’ defect reports. In: 2008 16th IEEE International Requirements Engineering Conference. IEEE, pp 135–144
- Hüttermann M (2012) Devops for developers. Apress
- ISO (2011) ISO/IEC 25010:2011(E): Systems and Software Engineering - Systems and Software Quality Requirements and Evaluation SquaRE
- Jha N, Mahmoud A (2019) Mining non-functional requirements from app store reviews. *Empirical Software Engineering*:1–37
- Kittlaus H-B, Fricker SA (2017) Software product management : the ISPM-compliant study guide and handbook. Springer
- Lu M, Liang P (2017) Automatic classification of non-functional requirements from augmented app user reviews. In: Proceedings of the 21st International Conference on Evaluation and Assessment in Software Engineering. ACM, pp 344–353
- Maalek W, Nayeibi M, Johann T, Ruhe G (2015) Toward data-driven requirements engineering. *IEEE Softw* 33(1):48–54
- Olsson T, Wnuk K (2018) QREME - Quality Requirements Management model for supporting decision-making. In: Requirements Engineering: Foundation for Software Quality
- Olsson T, Wnuk K, Gorschek T (2019) An empirical study on decision making for quality requirements. *J Syst Softw* 149:217–233
- Pettersson F, Ivarsson M, Gorschek T, Öhman P (2008) A practitioner’s guide to light weight software process assessment and improvement planning. *J Syst Softw* 81(6):972–995
- Regnell B, Brinkkemper S (2005) Market-driven requirements engineering for software products. In: Engineering and Managing Software Requirements, pp 287–308
- Rissanen O, Münch J (2015) Continuous experimentation in the b2b domain: a case study. In: Proceedings of the Second International Workshop on Rapid Continuous Software Engineering. IEEE Press, pp 12–18
- Runeson P, Höst M, Rainer A, Regnell B (2012) Case Study Research in Software Engineering
- Schwaber K (2004) Agile project management with scrum. Microsoft Press
- Seaman CB (1999) Qualitative methods in empirical studies of software engineering. *IEEE Trans Softw Eng* 25(4):557–572
- Seddon PB, Scheepers R (2012) Towards the improved treatment of generalization of knowledge claims in is research: drawing general conclusions from samples. *Eur J Inf Syst* 21(1):6–21
- Shahrokni A, Feldt R (2013) Industrial challenges with quality requirements in safety critical software systems. In: 2013 39th Euromicro Conference on Software Engineering and Advanced Applications. IEEE, pp 78–81
- Svahnberg M, Gorschek T, Nguyen TTL, Nguyen M (2015) Uni-REPM: a framework for requirements engineering process assessment. *Requir Eng* 20(1):91–118
- Sveningson R, Mattos DI, Bosch J (2019) Continuous experimentation for software organizations with low control of roadmap and a large distance to users: An exploratory case study. In: International Conference on Product-Focused Software Process Improvement. Springer, pp 528–544

- Wang T, Liang P, Lu M (2018) What aspects do non-functional requirements in app user reviews describe? an exploratory and comparative study. In: 2018 25th Asia-Pacific Software Engineering Conference (APSEC). IEEE, pp 494–503
- Wieringa R, Daneva M (2015) Six strategies for generalizing software engineering theories. *Sci Comput Programm* 101:136–152
- Wnuk K, Kollu RK (2016) A systematic mapping study on requirements scoping. In: Proceedings of the 20th International Conference on Evaluation and Assessment in Software Engineering - EASE '16, pp 1–11

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Thomas Olsson



Krzysztof Wnuk



Slinger Jansen



Affiliations

Thomas Olsson¹  · Krzysztof Wnuk² · Slinger Jansen^{3,4}

Krzysztof Wnuk
krzysztof.wnuk@bth.se

Slinger Jansen
slinger.jansen@uu.nl

¹ RISE Reserach Institutes of Sweden, Scheelevägen 17, 223 70 Lund, Sweden

² Software Engineering Research Group, Blekinge Institute of Technology, Karlskrona, Sweden

³ Utrecht University, Utrecht, The Netherlands

⁴ Visiting Scientist, School of Engineering Science, LUT University, Lappeenranta, Finland