2018 International Conference on the Quality of Information and Communications Technology

How Good Are My Requirements?

– a new perspective on the quality measurement of textual requirements –

Patrick Kummler
Karlsruhe Service Research Institute
Karlsruhe Institute of Technology
Karlsruhe, Germany
patrick.kummler@kit.edu

Léa Vernisse
Technical Unit ACM
Altran Deutschland S.A.S. & Co. KG
Munich, Germany
lea.vernisse2@altran.com

Hansjörg Fromm
Karlsruhe Service Research Institute
Karlsruhe Institute of Technology
Karlsruhe, Germany
hansjoerg.fromm@kit.edu

***Abstract*—Many efforts have been targeted towards securing and improving the quality of software requirements. Since the majority of software requirements today are still written in natural language, these approaches focus on measurable indicators that can be derived from the text. Our research particularly addresses the relationships between textual indicators and individual quality attributes, as worldwide industry standards have defined them. The study is conducted in connection with current development projects of the German automotive industry.**

Keywords*—*Requirements Quality, Automotive Industry, Software Requirements, Quality Attributes

I. Introduction

The digital transformation in the automotive industry faces manufacturers with new and unknown competitive challenges [1], [2]. "The global automotive industry is at the vanguard of a digital revolution" [3] and the value creation increasingly shifts from hardware to software and services [4].

Software requirements engineering is becoming more and more important for the automotive industry. One reason is that the role of software in the automobile generally changes: more than 90 percent of today's innovation comes from software and electronic systems [5]. Software and services have a significant influence on the customers' purchase decision and, accordingly, have become a main differentiator in the market.

At the same time, the complexity and criticality of software systems in the car increases with new applications like driver assistance, safety control, or autonomous driving. On the other hand, the automotive industry has become more networked with other industries than ever. Development is often done in collaboration with a number of partners (manufacturers, suppliers, engineering companies, mobility providers). This requires intense communication between the partners throughout the entire product lifecycle, but particularly in the early phases where the requirements are defined and communicated.

It has long been recognized that the success of a software development project strongly depends on the quality of the underlying requirements [6], [7], [8]. Failures in any stage of the product lifecycle can often be traced back to miscommunication in the early requirements phase. Appropriate methods, processes, and tools have been made available to secure the quality of software requirements. In many cases, the requirements quality is still at an insufficient level that results in additional effort for development and testing.

With the increasing volume, complexity, and criticality of software, new, reliable approaches are necessary to tackle the requirements quality problem. These approaches can be found in the areas of analytics and machine learning. The objective of our research is to identify such approaches that cope with the current challenges of the industry.

The research is in cooperation with an international automotive engineering and consulting company and enables us to have access on software requirements from the automotive industry. The available dataset consists of 32.735 requirements from different software projects to develop safety functions, comfort functions and digital automotive services.

II. Requirements Quality Analysis

Requirements quality is a well-established concept that is described in numerous text books and articles. Worldwide standards for requirements engineering are available which include sections on requirements quality. We refer to the standard ISO/IEC/IEEE 29148:2011 [9]. This standard contains a section on requirements characteristics that we use as a framework throughout our research.

It must be noted that the majority of software requirements today are still written in natural language and organized in text documents. Formal specification methods, known since decades, have hardly found their way into practice. Accordingly, most of the discussions on requirements quality focus on the quality of textual requirements. Reference [9], in consideration of [10], lists the following characteristics that requirements should possess: they should be *necessary*, *implementation free*, *clear and concise*, *consistent*, *complete*, *singular*, *feasible*, *traceable*, and *verifiable*. Other characteristics, also called attributes, have been proposed in the literature. For an overview, see e.g. [11].

Quality attributes are qualitative, they can only be judged and not be measured. Therefore, researchers have tried to identify indicators which can be measured quantitatively. Thus, an open research question is if these measurable indicators describe the individual quality attributes sufficiently.

A systematic analysis of the quality of requirements must necessarily concentrate on the requirement text itself or on measurable indicators that can be derived from the text. The literature on requirements quality analysis falls largely into three categories of research:

- The first category describes tools that are designed to assist the requirements engineer in writing good requirements. These tools identify weaknesses and give indications how the requirements can be improved in quality. Examples are RQA (Requirements Quality Analyzer) [11], ARM

978-1-5386-5841-3/18/$31.00 ©2018 IEEE
DOI 10.1109/QUATIC.2018.00031

(Automated Requirement Measurement) [12] and QuARS (Quality Analyzer for Requirements Specifications) [13], [14].

- A second category of research is devoted to the transformation of textual requirements into formal models and logic specifications. This is typically done by a thorough linguistic analysis of the requirements language. Examples are [15], [16], [17] and [18].
- A third category is aiming at the automated classification of a set of requirements according to good or bad overall quality. Either rule-based or machine learning methods are used. Examples are [19], [20], [21] and [22].

Most of the papers, but not all, know the concept of quality attributes. Some papers address a complete set of quality attributes [11], [12], [23], other papers focus on selected attributes like ambiguity [19], [24], [25], readability [17], atomicity [20], completeness and consistence [17], [18], certainty [21], and correctness [22].

All papers describe textual indicators that are believed to relate to the quality of the requirements. These indicators can be lexical, syntactic, or semantic [12], [26]. Some papers suppose a direct relationship between the indicators and the overall requirement quality.

Reference [19] uses 12 textual indicators fed into a decision tree to classify a requirement into ambiguous or non-ambiguous. Reference [22] lists desirable properties of a *good* requirement and focusses on metrics for correctness. Their method is composed of two tasks. In the first task, they generate classifiers. They let experts assess requirements according to their quality and use 24 textual indicators fed into a decision tree to classify requirements into good or bad. In the second task, they estimate and evaluate the quality of new requirements based on the same set of metrics. Both papers compare the performance of the machine learning algorithm (decision tree) against expert judgements (ambiguous/non-ambiguous, good/bad) on the individual requirements.

Other papers suggest relationships between textual indicators and individual quality attributes [11], [12], [23]. In no case, these relationships are proven by empirical research. In the next chapter, a research approach is presented which provides a structured method for our intention to close this research gap and to reveal relations between indicators and quality attributes.

III. Research Approach

In order to study the relationship between textual indicators and individual quality attributes, the following steps are necessary:

- 1st step: Definition of requirements quality, related quality attributes and methods for quality measurements
- 2nd step: Derivation of indicators from textual requirements using text mining and natural language processing methods
- 3rd step: Judgement on individual quality attributes of a requirement by a group of experts
- 4th step: Formulation of a regression model to estimate the relationship between indicators and attributes

This analysis will be complemented by the following steps:

- 5th step: Judgement on the overall quality of a requirement by the group of experts
- 6th step: Formulation of a regression model to estimate the relationship between individual quality attributes and overall quality of a requirement

To conduct this study, it must be recognized that the requirements quality attributes as defined by [9] and [10] are of quite different nature. We can distinguish attributes whose fulfillment for a requirement can be answered with yes or no, and others, whose fulfillment can be answered with more or less. Examples of the first kind are *consistent*, *complete*, or *singular*. Examples of the second kind are *necessary* or *feasible*. Some attributes can be judged by looking at a requirement in isolation (e. g. singularity), others need a look into the entire requirements document (e. g. traceability).

Most remarkable, however, is the special role that *clear and concise* has over all other attributes. If a requirement is not *clear and concise*, there is reason for misunderstanding and misinterpretation. But, moreover, if a requirement is not *clear and concise*, it can hardly be assessed to which degree it fulfills the other attributes. This is an assumption that other authors describe with “surface understanding” vs. “concept understanding” [19], or “clarity” vs. “content” [27]. The clarity or surface level is concerned with the language of the requirement, or “what is stated” [19], and the concept or content level with “what is meant or implied” [19]. It is clear that an assessment of the latter requires a much deeper view into the structure of the requirement – for a human evaluator as well as for a machine algorithm. Accordingly, analyses on the content level fall into our second category of research: they are based on the transformation of textual requirements into formal models and logic specifications in order to detect e.g. consistency or completeness problems. Fig. 1 shows the relation between *clear and concise* and the other quality characteristics. The verification of this relation is part of our future research.

The surface level has been most often addressed by research. Especially the tools that have been developed to support the requirements engineer in writing good requirements (our first research category) work largely on the surface level.

The textual features that have been suggested in literature can widely be associated with the surface level. Yet, some researchers assume an influence of these indicators on content properties. The hypothesis might be that an author, who puts much effort in writing a requirement *clearly and concisely* from a textual point of view, has put the same effort in assuring that the requirement is *complete*, *consistent*, *verifiable*, etc.

Fig. 2 depicts this situation. For illustration purposes, the indicators used in [12] have been selected as the independent variables on the left. The quality attributes are the ones described in [9] and [10], which are all believed to contribute to the overall requirement quality. Assumed influences between indicators and attributes have exemplarily been

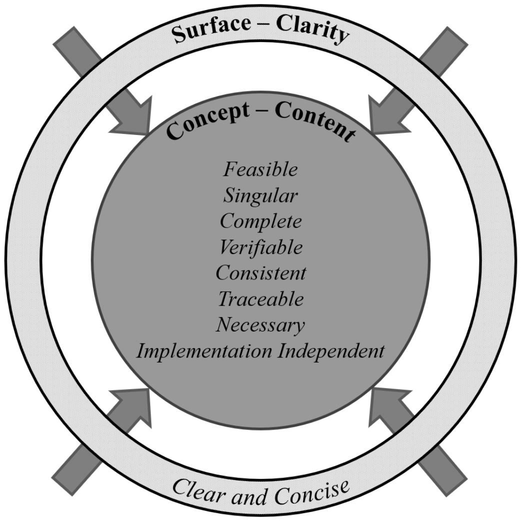

Fig. 2. If the requirement is not *clear and concise*, it is blurring the view on other attributes

taken from [12] for the three quality attributes *clear and concise*, *complete*, and *consistent*.

The evaluation of requirements through experts enables us to link between textual indicators found in the literature and quality attributes promoted by experts. Working closely with the major automotive companies allows us to work within an industrial environment and in direct contact with their experts. Our intention is to automatically extract textual features from a large set of requirements with text mining approaches. This includes lexical and syntactical analysis accompanied by semantic knowledge (vocabularies, glossaries).

For this we intend to conduct a survey among experts, who will be asked to judge on each individual attribute and the overall quality of each requirement. In a first regression model, the values of the indicators will be related to the attribute ratings of the experts. The regression will show if the influences conjectured by several authors can be confirmed. In a second regression model, the individual attribute ratings will be related to the overall quality rating.

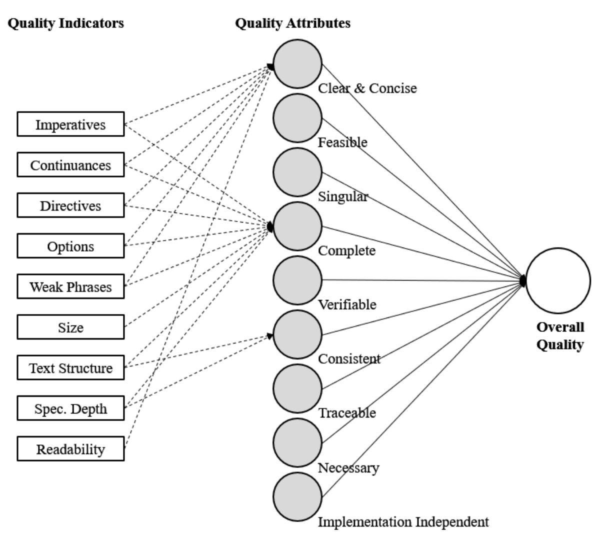

Fig. 1. Hypothesis model of quality indicators and attributes

This will show which attributes are most influential on the overall quality, and, at the same time, reveal mutual influences between the attributes, as it has been suggested by [11].

As a preliminary step, we already asked requirement experts from the automotive sector about the perceived importance of the individual quality attributes (1 = not important, 9 = very important). The experts have distinct roles: requirements manager and engineer, software architect and tester, failure manager and system integrator with an average five years of professional experience. Table I shows the aggregated results with average, median and variance for each attribute. Some of the attributes are evaluated with very small variance among the experts: *implementation independent* with 0.54, *consistent* with 0.87, *traceable* with 0.98 and *clear and concise* with 1.04. This implies a mutual understanding about the importance of several attributes between the interviewees. The experts nominate *clear and concise* as the most important attribute. The remaining attributes have a comparable high variance value between 3.20 for *complete* and 4.89 for *feasible*.

The importance of *clear and concise* can certainly be explained with the special role (as described above) that this attribute has compared to the others.

The high variance of *complete*, *feasible*, *singular*, *verifiable*, and *necessary* could either be explained by the different background and experience of the experts or by the fact that these attributes themselves are not always very easy to understand. A deeper analysis of this survey is part of our ongoing research project.

IV. CONCLUSION

This paper sets up to analyze approaches for quality measurement of textual requirements. We identify the ISO/IEC/IEEE 29148:2011 standard as relevant source. The standard describes how requirements quality is defined and offers relevant attributes to measure requirements quality.

Our study has shown that even if a lot of research has been targeted towards the textual quality of software requirements, relationships between textual indicators and individual quality attributes have been suggested, but not evidenced so far.

We outlined how we plan to close this research gap. Indicators will be derived from textual requirements using text mining and natural language processing methods. The individual quality attributes of requirements will be judged by a group of experts.

TABLE I. PRIORITIZED QUALITY CHARACTERISTICS FROM EXPERTS.

	Average	Median	Variance
clear and concise	7.42	7.50	1.04
feasible	7.33	8.50	4.89
singular	6.75	7.50	3.65
complete	6.42	6.75	3.20
verifiable	5.17	4.00	4.14
consistent	4.92	5.00	0.87
traceable	2.75	3.00	0.98
necessary	2.50	1.00	4.58
implementation independent	2.08	2.00	0.54

A regression model will be built that estimates the relationship between the indicators and attributes. Additionally, the group of experts will judge on the overall quality of the requirements, so that the relationship between individual quality attributes and overall quality of a requirement can be analyzed as well. In this attempt, one has to be aware of the quite different nature of the established quality attributes. We have identified that *clear and concise* assumes a special role among the attributes. This is also reflected by the first rank that *clear and concise* has received in our expert survey on importance. The variance of answers regarding the other attributes, however, shows that these attributes are seen inconsistently by the experts. We will have a deeper look into this phenomenon in our further research.

REFERENCES

[1] A. Hanelt, E. Piccinini, R.W. Gregory, B. Hildebrandt, and M. Lutz, “Digital transformation of primarily physical industries – exploring the impact of digital trends on business models of automobile manufacturers,” 12th International Conference on Wirtschaftsinformatik, pp. 1313–1327, 2015.

[2] S.J. Berman and R. Bell, “Digital transformation: creating new business models where digital meets physical,” IBM Institute for Business Value, 2011.

[3] J. Duncan, S. Lulla, A. Marshall, and B. Stanley, “Driving digital destiny - digital reinvention in automotive,” IBM Institute for Business Value, 2015.

[4] W. Boston, The future is electric for BMW - German luxury-car maker has announced a strategy shift, URL: https://www.wsj.com/articles/the-future-is-electric-for-bmw-1458142897 (visited on 20/04/2018).

[5] E. Hirsh, A. Kakkar, A. Singh, and R. Wilk, 2015 Auto Industry trends, URL: https://www.strategyand.pwc.com/reports/automotive-trends-2015 (visited on 20/04/2018).

[6] B. Curtis, H. Krasner, and N. Iscoe, “A field study of the software design process for large systems,” Communications of the ACM, vol. 31(11), pp. 1268–1287, 1988.

[7] M.I. Kamata and T. Tamai, “How does requirements quality relate to project success or failure?,” in Requirements Engineering Conference 2007, pp. 69–78, 2007.

[8] E. Knauss, C. El Boustani, and T. Flohr, “Investigating the impact of software requirements specification quality on project success,” in PROFES, pp. 28–42, 2009.

[9] ISO, “ISO/IEC/IEEE 29148:2011, Systems and software engineering - Life cycle processes - Requirements engineering,” Geneva, Switzerland, 2011.

[10] G. Fanmuy, J. Marvin, and H. Ronald, “Deployment Package - Needs and Requirements Engineering - Systems Engineering Basic Profile,” International Council on Systems Engineering (INCOSE), 2014.

[11] G. Génova, J.M. Fuentes, J. Llorens, O. Hurtado, and V. Moreno, “A framework to measure and improve the quality of textual requirements,” in Requirements Engineering, vol. 18(1), pp. 25–41, 2013.

[12] W.M. Wilson, L.H. Rosenberg, and L.E. Hyatt, “Automated analysis of requirement specifications,” in Proceedings - International Conference on Software Engineering, pp. 161–171, 1997.

[13] A. Fantechi, S. Gnesi, G. Lami, and A. Maccari, “Applications of linguistic techniques for use case analysis,” in Requirements Engineering, vol. 8(3), pp. 161–170, 2003.

[14] F. Fabbrini, M. Fusani, S. Gnesi, and G. Lami, “The linguistic approach to the natural language requirements quality: benefit of the use of an automatic tool,” in Proceeding of Software Engineering Workshop, 26th Annual NASA Goddard, pp. 97–105, 2001.

[15] M.G. Ilieva and O. Ormandjieva, “Automatic transition of natural language software requirements specification into formal presentation,” in International Conference on Application of Natural Language to Information Systems, pp. 392–397, June 2005.

[16] K. Verma and A. Kass, “Requirements analysis tool: A tool for automatically analyzing software requirements documents,” in International Semantic Web Conference, pp. 751–763, October 2008.

[17] J. Holtmann, J. Meyer, and M. von Detten, “Automatic validation and correction of formalized, textual requirements,” in Software Testing, Verification and Validation Workshops (ICSTW), pp. 486–495, March 2011.

[18] A. Arellano, E. Zontek-Carney, and M.A. Austin, “Frameworks for natural language processing of texual requirements,” in International Journal On Advances in Systems and Measurements, vol. 8, pp. 230–240, 2015.

[19] O. Ormandjieva, I. Hussain, and L. Kosseim, “Toward a text classification system for the quality assessment of software requirements written in natural language,” in Fourth international workshop on Software quality assurance: in conjunction with the 6th ESEC/FSE joint meeting, pp. 39–45, September 2007.

[20] C. Huertas and R. Juárez-Ramírez, “Towards assessing the quality of functional requirements using English/Spanish controlled languages and context free grammar,” in The Third International Conference on Digital Information and Communication Technology and its Applications (DICTAP2013), pp. 234–241, 2013.

[21] H. Yang, A. De Roeck, V. Gervasi, A. Willis, and B. Nuseibeh, “Speculative requirements: Automatic detection of uncertainty in natural language requirements,” in Requirements Engineering Conference (RE), 20th IEEE International, pp. 11–20, September 2012.

[22] E. Parra, C. Dimou, J. Llorens, V. Moreno, and A. Fraga, “A methodology for the classification of quality of requirements using machine learning techniques.” Information and Software Technology 67, pp. 180–195, 2015.

[23] E. Gallego et al., “Requirements quality analysis: A successful case study in the industry,” in Complex Systems Design & Management, pp. 187–201, 2017.

[24] D.N. Berry, “Ambiguity in natural language requirements documents,” in Monterey Workshop, pp. 1–7, September 2007.

[25] U.S. Shah and D.C. Jinwala, “Resolving ambiguities in natural language software requirements: a comprehensive survey,” ACM SIGSOFT Software Engineering Notes, vol. 40(5), pp. 1–7, 2015.

[26] V. Antinyan and M. Staron, “Rendex: A method for automated reviews of textual requirements,” in Journal of Systems and Software, vol. 131, pp. 63–77, 2017.

[27] A.J. Tyznik, S. Verma, Q. Wang, M. Kronenberg, and C.A. Benedict, “Distinct requirements for activation of NKT and NK cells during viral infection,” in The Journal of Immunology, vol. 192(8), pp. 3676–3685, 2014.