

Automatic Requirements Reviews - Potentials, Limitations and Practical Tool Support

Henning Femmer(✉)

Qualicen GmbH, Garching bei München, Germany
henning.femmer@qualicen.de

Abstract. Requirements are usually documented, and natural language is still the primary choice of syntax. However, in particular with natural language, the quality of the documentation is a key success factor for projects. To keep this risk in check, projects apply manual quality assurance in the form of reviews. Due to the shortcomings of manual reviews, more and more companies look into lightweight automatic support mechanisms to improve the quality of requirements documents.

1 Description

To document requirements, natural language is still the primary means. Requirements in natural language can be created and understood by all stakeholders without additional effort and specific requirements engineering background. However, natural language poses the risk of being imprecise or ambiguous. Badly written requirements have an expensive impact on the whole project. Incomplete or ambiguous requirements generate additional effort due to unnecessary feedback loops. In the end, bad requirements lead to misinterpretations and finally to the wrong product.

Manual reviews are an effective tool to create high quality requirements documents. Although effective, this method comes with considerable effort. The manual inspection of the requirements by multiple reviewers and the integration of review results are time consuming. As one review cycle often takes days or weeks to complete, the author of the requirements has to wait a long time before receiving feedback. The result of these problems is that reviews are often only performed sporadically or only superficially.

2 Potentials

The potential for automatic reviews detection lays in the aforementioned deficiencies of manual reviews; A claim that is also supported by practitioners and evaluated in practice (see, e.g. [5]).

Automatic reviews are cheap. As we discuss in [3], one of the key challenges to establishing manual reviews in practice are the high costs that come with a thorough analysis. Therefore, to have a mechanism that provides feedback free of charge is a promising advantage. (For the sake of the argument, we deliberately ignore the costs of setting up and maintaining such an analysis.)

Automatic reviews are fast. For many quality aspects, feedback can be given more or less instantaneously. For example, in the current configuration, our requirements scout receives, processes, and renders feedback for a paragraph in around 500 ms. This enables us to give feedback on-the-fly.

Automatic reviews are consistent. If you hand a requirements artifact to one reviewer on two different days, you might receive two very different results. The review process depends on various personal factors, such as the state of mind or the recent input of the reviewer. While this can be an advantage, for quality factors such as *Does the artifact follow a guideline?*, this subjectivity throws the doors wide open for inconsistencies. An automatic method works in a deterministic way, every day.

In an analysis of an industrial RE artifact guideline in [6], we estimate that 52% of the guideline rules can be automatically checked perfectly or with a good heuristic. Surprisingly, for detection of quality violations, most rules require just simple heuristics.

Accordingly, many researchers work on providing automatic support for REs [1, 4, 5, 7–10].

3 Limitations

Various reasons imply that a quality factor cannot be automatically detected. This is analyzed in-depth, i.a. in [2] or [6]. The main reasons, why certain quality factor cannot be automatically detected are:

- Automatic reviews require an explicit quality definition.
- Automatic reviews struggle with noise in industrial data.
- Automatic reviews have no deep semantic understanding of text.
- Automatic reviews have no knowledge of domain and common sense.
- Automatic reviews do not know the goal of the system and the current project status.

4 Tooling: The Requirements Scout

Tooling for automatic requirements reviews can be seen from three roles: The central role is, of course, the requirements author, who produces the requirements. But there is also the reviewer, who proof-reads and validates the requirements. And finally, there is the QA-Engineer, responsible for the overall quality of all artifacts created during the engineering process. Each of these roles needs a different view on requirements and different tools in order to do their work efficiently and achieve a high requirements quality.

The author is interested in direct feedback. In a previous analysis [3], we found that receiving feedback directly within the tool increases the willingness of authors to use such a tool. Therefore, the requirements scout comes with various plugins, e.g. for the widespread RE tool PTC Integrity (see Fig. 1) or Microsoft Word.

Requirement Document...cation document (162)				
Show items containing <input type="text" value="where"/>				
ID	Section	Text	QRC_Findings	QRC_Findings_Text
178	2.4.2	Because of base waterproof protection. Watch face be made out of glass.	Passive Voice Smell	Because of base waterproof protection. Watch face be made out of glass.
188	2.5	Time Zone Requirements		Time Zone Requirements
190	2.5.1	All time zones must run of the same internal quartz crystal. To distinguish between the time zones, the embedded microprocessor must calculate the offset of each time zone and keep track of the time zones chosen by the user. yes		All time zones must run of the same internal quartz crystal. To distinguish between the time zones, the embedded microprocessor must calculate the offset of each time zone and keep track of the time zones chosen by the user. yes
192	2.5.1.1	As soon as one time zone time is changed, all others must adjust automatically. The only exception to this is a daylight savings time change. By knowing the offset of each time zone, the MPU can determine the new times for the other times zones using the appropriate offsets	Passive Voice Smell Imprecise Phrase Smell	As soon as one time zone time is changed, all others must adjust automatically. The only exception to this is a daylight savings time change. By knowing the offset of each time zone, the MPU can determine the new times for the other times zones using the appropriate offsets
194	2.6	Timer Requirements		Timer Requirements
196	2.6.1	tbd.		tbd.
198	2.6.1.1	For the timer feature, the counter on the digital display must be counted down to 00:00:00.	Passive Voice Smell	For the timer feature, the counter on the digital display must be counted down to 00:00:00.
200	2.7	Chronometer Requirements		Chronometer Requirements
202	2.7.1	The chronometer must have an easy-to-understand user interface which allows the user to quickly clear and restart the chronometer or restart timing from their previous stopped time.	Long Sentence Smell	The chronometer must have an easy-to-understand user interface which allows the user to quickly clear and restart the chronometer or restart timing from their previous stopped time.

Fig. 1. PTC integrity integration of qualicen scout

The reviewer is interested in focussing on the most relevant quality factors. For example, he or she wants to analyze whether the specification validly meets the stakeholders' goals. Therefore, the scout offers the option to list all findings that the scout detects, so that the reviewer quickly picks the interesting ones from the list and then continues focussing on the content.

Lastly, the QA-Engineer (and also management) is interested in the development of quality over time. Therefore, the scout enables to set baselines, compare two versions, and analyze trends and tendencies in order to understand in which direction the quality is going (see Fig. 2).

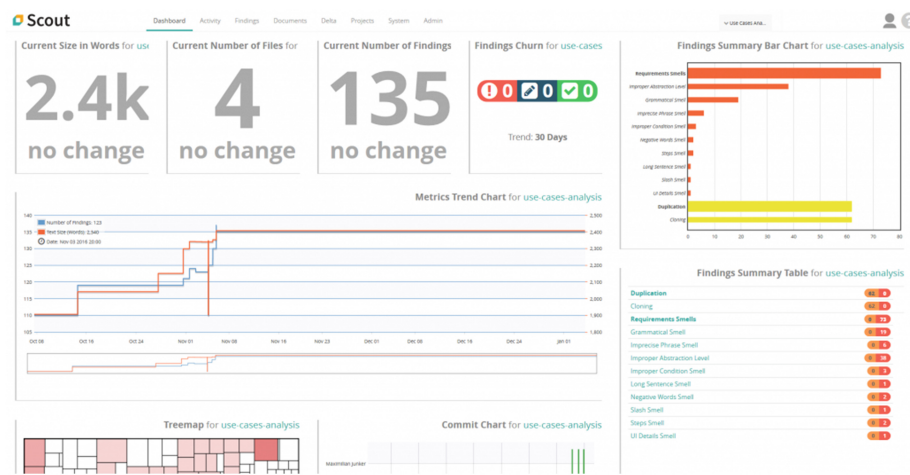


Fig. 2. Qualicen scout quality dashboard

5 Summary

Automatic review techniques have matured over the last years: A substantial set of widespread quality defects in requirements documents can now be found automatically. Examples of such defects are ambiguous wording or overly complex sentences. Also more complex defects, such as cloning, inadequate levels of abstraction, or wrong references in documents can be detected automatically.

However, there are also strong limitations to automatic review techniques. In a recent study [6] we quantified this, estimating that 52% of the guideline rules can be automatically checked perfectly or with a good heuristic. As a conclusion, automatic reviews do not replace, but complement manual reviews. It is assumed, however still remains to be shown, that automatic reviews reduce the time needed for manual reviews and provide faster and less expensive feedback for requirements authors.

References

1. Fabrini, F., Fusani, M., Gnesi, S., Lami, G.: Quality evaluation of software requirements. In: Software and Internet Quality Week Conference, pp. 1–18 (2000)
2. Femmer, H.: Requirements engineering artifact quality: definition and control. Ph.D. thesis, Technische Universitat Munchen (2017)
3. Femmer, H., Hauptmann, B., Eder, S., Moser, D.: Quality assurance of requirements artifacts in practice: a case study and a process proposal. In: Abrahamsson, P., Jedlitschka, A., Nguyen Duc, A., Felderer, M., Amasaki, S., Mikkonen, T. (eds.) PROFES 2016. LNCS, vol. 10027, pp. 506–516. Springer, Cham (2016). doi:[10.1007/978-3-319-49094-6_36](https://doi.org/10.1007/978-3-319-49094-6_36)
4. Femmer, H., Mendez Fernandez, D., Juergens, E., Klose, M., Zimmer, I., Zimmer, J.: Rapid requirements checks with requirements smells: two case studies. In: RCoSE, pp. 10–19. ACM (2014)
5. Femmer, H., Mendez Fernandez, D., Wagner, S., Eder, S.: Rapid quality assurance with requirements smells. *J. Syst. Soft.* **123**, 190–213 (2017)
6. Femmer, H., Unterkalmsteiner, M., Gorschek, T.: Which requirements artifact quality defects are automatically detectable? A case study. In: AIRE, pp. 1–7. IEEE (2017)
7. Juergens, E., Deissenboeck, F., Feilkas, M., Hummel, B., Schaetz, B., Wagner, S., Domann, C., Streit, J.: Can clone detection support quality assessments of requirements specifications? In: ICSE. ACM (2010)
8. Krisch, J., Houdek, F.: The myth of bad passive voice and weak words: an empirical investigation in the automotive industry. In: RE. IEEE (2015)
9. Lucassen, G., Dalpiaz, F., van der Werf, J.M.E.M., Brinkkemper, S.: Improving agile requirements: the quality user story framework and tool. *Requirements Eng. J.* **21**(3), 383–403 (2016)
10. Wilson, W.M., Rosenberg, L.H., Hyatt, L.E.: Automated analysis of requirement specifications. In: ICSE, pp. 161–171. ACM (1997)