

**Jerzy Kocerka**  
(imię i nazwisko kandydata)

**Gliwice, 7.02.2023**  
(miejsowość, data)

[REDACTED]  
[REDACTED]  
(adres do korespondencji)

[REDACTED]  
[REDACTED]  
(nr telefonu i adres e-mail)

**Thesis Title:** Automatic quality analysis of textual requirements using natural language processing

### **Abstract**

The doctoral thesis examines the utilization of Natural Language Processing (NLP) for the automatic analysis of the quality of industrial software requirements and the impact of requirement quality on software production efficiency.

The study reviews and analyzes various forms of requirement documentation, including formal, semi-formal and natural language methods. The research also verifies different quality metrics for natural language requirements, including the presence of grammatical and spelling errors, subjective and weak definitions, the absence of references to other requirements or documents, inadequate requirement structure, and various types of ambiguity. Existing algorithms and requirement analysis tools based on machine learning and other techniques (e.g., dictionary-based) are also discussed.

It also presents the current state of the possibilities regarding tools for natural language processing, in particular for text analysis and information extraction, based on artificial intelligence algorithms.

Based on the analyzed metrics, an original quality model for natural language requirements was created and served as the foundation for the development of an NLP-based tool for automatic requirement analysis. The tool performs analysis and generates various requirement metrics, and then calculates an overall quality score for each requirement by analyzing key quality metrics.

A set of requirements with expert annotations was used to create a reference for automatic analysis. The results obtained from the automatic tool were compared with expert annotations to gain insight into the strengths and weaknesses of NLP-based requirement analysis.

To verify the thesis hypothesis that automatic requirement quality analysis improves the efficiency of software creation and testing for embedded systems, a comparison was made between the requirement creation and review processes with and without the use of the tool.

The thesis concludes with a summary and description of the potential for implementing the proposed solution and directions for future research.

**Keywords:** Requirements quality, natural language processing, software development