Assignment 3 - The Article

Student: av222vu – Course: 1DV609

Article: https://onlinelibrary-wiley-com.proxy.lnu.se/doi/epdf/10.1002/smr.2741

1. Why and what problem are the author(s) trying to solve?

The authors address the challenge of ensuring software quality and reliability in AI-augmented software engineering. As generative AI is increasingly being integrated into the software development lifecycle, it introduces risks of bias, 'hallucinations' (the potential for making incorrect inferences), lack of transparency and inconsistent quality.

The key concern is that existing ISO/IEC (Information Security Standard) AI quality models are not sufficient in fully evaluating AI systems. These models focus narrowly on traditional software quality without reflecting on AI-specific attributes like explainability, ethical compliance and resilience to manipulation.

The problem lies in the lack of frameworks that address:
– Trust and safety in AI systems.
– User alignment and fairness.
– Ethical boundaries and governance.
– Reliability across complex real world use cases.

2. How and what method has been used to answer the question?

The methods used in the article were:

Literature review and model analysis:
They analysed industry standards such as the ISO/IEC 25059 and ISO/IEC 25010 models, which guide software and AI quality assessments. Which they used to identify gaps in AI-specific standards.

Industry Survey:
A survey was done with over 150 industry practitioners and academics. Participants were asked,
"What does good quality look like in AI and machine learning systems (in one or two words)?"
The responses were displayed using the following word cloud to highlight recurring words.

Model Enhancement:

Based on the survey, the authors proposed 45 additional sub-characteristics for the ISO/IEC 25059 framework, addressing areas such as ethics, empathy, resilience and explainability.

3. What is the result from the method?

The survey revealed a consensus on key AI quality traits such as trustworthiness, fairness, interpretability, robustness, empathy, and non-corruptibility. AI systems were expected to display human-like qualities (e.g., ethical behavior, emotional intelligence).

From the results the authors proposed two extended models:
ISO/IEC 25059 Product Quality Model (45 new sub-characteristics).
ISO/IEC 25059 Quality in Use Model (additional quality-in-use traits).

Enhancements include characteristics related to security, societal impact, and AI transparency. This is being illustrated through new, extended diagram models with new sub-characteristics and aligning them with AI governance and trustworthiness frameworks (EU AI Ethics Guidelines).

4. What is the validity of the result according to the author?

The authors recognize that while their proposed models and characteristics reflect industry feedback, their validity is still in the early stages and needs further testing to be done.

One of the strengths of the study is that the survey captures real-world insights from professionals who work directly with AI systems. Additionally, the proposed

extensions build on well-established ISO/IEC frameworks, which helps give them credibility. However, the validity of the findings comes with some limitations.

– The survey included 150 industry practitioners and academics, which may not fully represent all AI applications or industries.
– No formal pilot study has been conducted to test the models in practice.
– There is also a risk that the findings may not apply equally across different AI fields, such as healthcare AI compared to autonomous vehicles.

Limitations of Validity:
- The survey sample size (150) may not fully represent all AI use cases or industries.
- No formal pilot study of the proposed model is conducted yet.
- Generalization risks exist due to the diversity of AI applications (e.g., healthcare AI vs. autonomous vehicles).

5. Does the research solve the problem/answer the question?

The paper successfully highlights gaps in existing AI quality models and provides a concrete, survey-backed proposal for improvement. However the effectiveness of the proposed extension is not yet tested in real world software development environments.
While the research promotes the conversation, further empirical testing and refinement are necessary to ensure that these model extensions cover AI system complexities.

6. What is to be done next, future work, and why?

Model Testing and Validation

The authors plan to pilot the enhanced models across various industries to measure their effectiveness. Real-world case studies will help refine the sub-characteristics and validate their impact on AI reliability and quality.

Alignment with Emerging AI Standards

Future research will explore other AI-related ISO standards, such as:

ISO/IEC TR 24027 (Bias in AI Systems)
ISO/IEC TR 24028 (Trustworthiness in AI)
EU AI Act guidelines

Academic Collaboration

The authors propose integrating their findings into the university curricula, enhancing AI ethics and quality engineering education.

Future work is needed because without formal validation, there is a risk of oversimplification or underestimating complex AI risks. Testing is crucial to ensure the proposed characteristics effectively mitigate real-world AI failures.

7. What do you think and why?

I believe this paper addresses a gap in AI quality assurance by expanding existing models to reflect AI-specific risks and attributes. The reliance on survey data and industry input strengthens the credibility of the proposed extensions.
However, while the paper highlights practical concerns (e.g., bias, ethics), it lacks quantitative analysis of AI failures and their impact, a limitation the authors themselves acknowledge. [6, Summary]
Real case studies of AI-related failures could have enriched the discussion.

Moreover, addressing emotional and ethical characteristics in AI quality models represents a progressive shift, but it raises other questions:
– How do we measure empathy in AI?
– Can emotional intelligence be tested?