Operationalizing Ethical Principles in AI for Healthcare through Agile Software Development Lifecycle

Lameck Mbangula Amugongo [1,2]

[1]Namibia University of Science & technology   [2]IndabaX Namibia

Introduction

The growing application of machine learning (ML) techniques in healthcare has increased awareness of the ethical issues that arise in the design, deployment and use of AI systems. Ethical issues such as privacy, accountability, transparency, fairness, robustness, safety and trust have been widely reported and discussed in the literature, and if not considered, ethical concerns can pose a threat to equitable health delivery and human rights [1].

Critical questions have been raised about whether ethical principles can significantly influence the decision-making processes of humans working in the field of AI and ML because AI ethics lack the means to reinforce its own normative beliefs.

Existing approaches

A framework to operationalise ethics based on existing guidelines that provide actionable solutions [2].

Problem statement

AI development life cycle mostly follows the software development life cycle (SDLC). Therefore, it would make more sense for ethical principles to Incorporated in different phases of the SDLC.

Methodology

Our framework is conceived and applied using the "ethics by design" approach, where ethical principles are incorporated iterative throughout the entire AI development pipeline (from requirements elicitation to deployment and maintenance) in the agile SDLC. Therefore, our approach ensures that ethical considerations are not just an afterthought but practically integrated into the development of AI solutions.

Ethical principles for AI-powered systems for Healthcare

Our ethical values encompass relational and communal aspects because we believe that healthcare is intrinsically a matter that affects all of society. Moreover, we posit that the relational aspects will ground AI systems for healthcare to be informed by the lived experiences of all patients, particularly those who are disproportionally affected by algorithmic injustices. We identified seven (7) ethical principles:

- **Fairness:** entails equal treatment, no discrimination, equity and empowerment. Focusing on the disproportionately impacted, i.e. the most marginalised and underrepresented communities [3].
- **Agility:** refers to the ability of the algorithm to capture temporal changes in clinical events that occurs.
- **Precision:** ability of the model to accurately perform prediction tasks.
- **Safeguarding humanity:** imply that the developers or other crucial decision-makers are considering potential harm to patients and the community as a whole.
- **Respect for others:** exhibit compassion, solidarity and care for all users especially the most vulnerable.
- **Trust and accountability:** trust is more than just technical aspects. Need to establishing long-term relationships between users and AI systems.
- **Robust and reproducibility:** ensure that the AI system operates reliably throughout its entire lifecycle.

Operationalizing ethical principles into agile SDLC

Figure 1. illustrates how the identified ethical values and principles in section 3 can be practically incorporated into the agile SDLC. The principle we have identified should be addressed in different phases of the SDLC throughout the AI lifecycle.
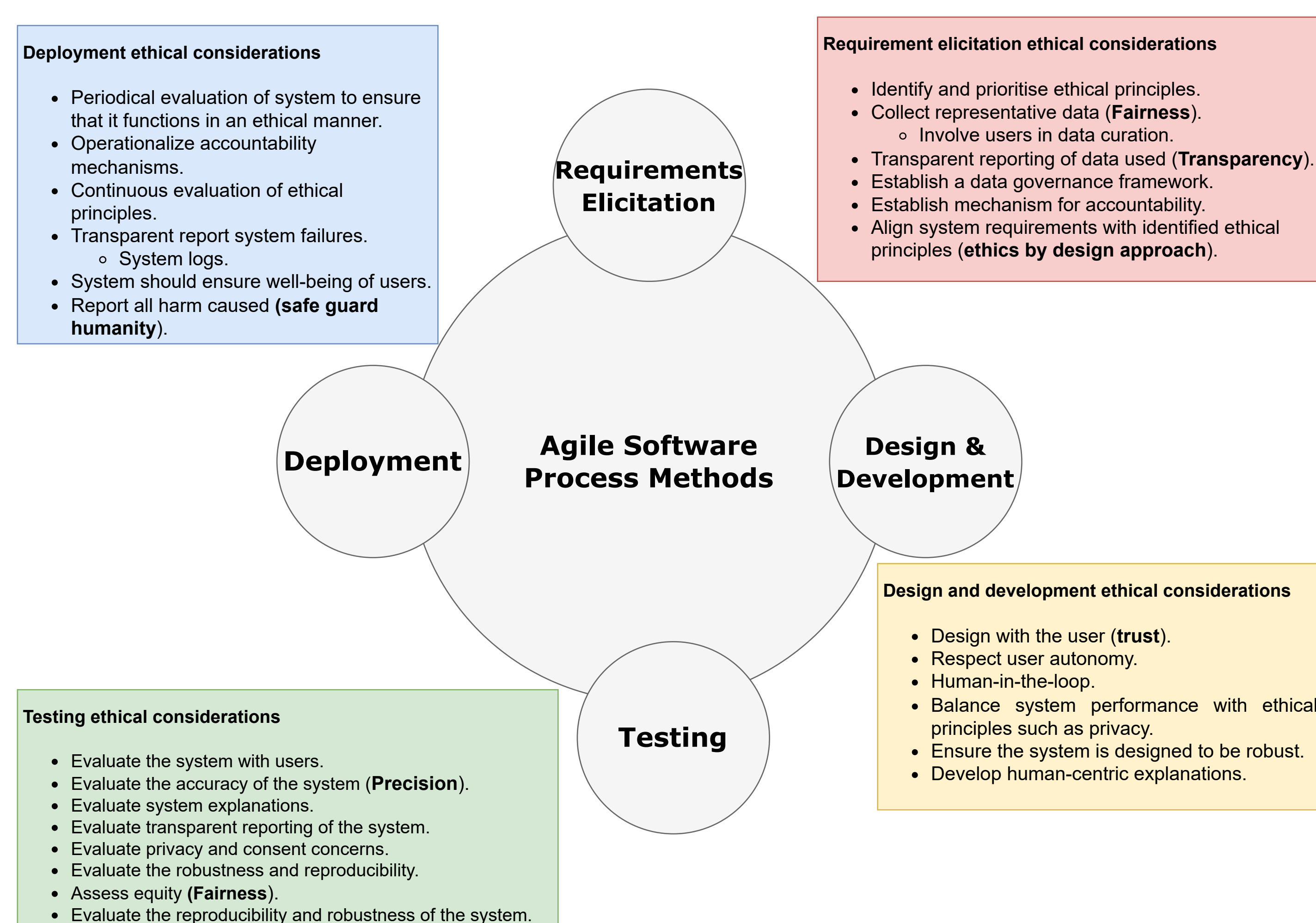

Figure 1. Embedding ethical principles into the agile software development process.

Ethics by Design: Implementing AI ethics through SDLC phases

Creating AI systems that meet normative standards is a challenging task that cannot be solved simply by urging developers to be more "ethical" in their work. Due to the complex nature of AI systems, we cannot know all of their outcomes in advance. We provide contextualized ethical principles most relevant to the particular phase in the SDLC.

Requirement Elicitation

- Identify ethical principles and align with system requirements
- Establish data governance framework
- Curate representative dataset
- User involvement in key decision from onset

Design and development

- Design with user
- User autonomy
- Balance system performance with privacy
- Ensure robust system design
- Develop human-centric explanations

Testing

- Evaluate system with the user
- Evaluate AI explanations
- Evalaute privacy and consent concerns
- Assess equity
- Assess group and individual fairness
- Evaluate system reproducibility and robustness

Deployment

- Operationalise accountability mechanisms
- Periodic system evaluation
- Transparent reporting of system failures
- Document and report harm causer by systems

Discussion

Though helpful, codes of ethics and conduct alone are insufficient for effectively addressing AI systems development values and principles. Therefore, comprehensive measures are necessary such as operationalizing AI ethical values and principles throughout the agile AI system development lifecycle. The agile development lifecycle comprises four phases: requirements elicitation, design and development, testing, and deployment.

Moreover, we recommend that development teams consider techniques to assess how the developed AI system aligns with the 12 principles of agile software development, which include (1) user satisfaction; (2) accommodating requirements changes; (3) frequent delivery of working software; (4) collaboration with stakeholders; (5) support, trust and motivation; (6) Effective communication; (7) Measure progress through functional delivery of system; (8) Consistency in development and maintenance; (9) Attention to technical detail and improve design agility; (10) Simplicity; (11) Self-organising teams; (12) Regular reflections on how to improve effectiveness in the team [4].

Limitations

1. Acknowledge that we lack oversight mechanisms to effectively align AI development in the healthcare sector. We believe that regulations and governance mechanisms play a crucial role in ensuring the alignment of AI with ethical values.

2. Provide a practical demonstration of how ethics can be integrated into the agile software development process using a specific use case. Further studies are needed to showcase how AI ethical principles can be operationalized to other healthcare problems, such as medical imaging.

3. Understand that there may be differences between the conceptual and practical implications of our proposed framework. Thus, we provide an unambiguous way to operationalise ethical principles for AI systems for healthcare within the agile process which developers of mobile applications are familiar with and trained to use.

4. Demonstrate how our framework can be operationalised using AI healthcare examples. Though the identified principles and guidelines are specific to AI-enabled mobile health solutions.

Our approach to operationalizing AI ethics in the SDLC throughout the AI development pipeline can be applied to other domains.

Conclusion

We propose a contextualized case-based framework that empowers developers to operationalize ethical principles within the agile. We emphasize the crucial role of community involvement in the development of human-centred AI systems for healthcare, advocating for co-designing AI systems with the local community and expert clinicians.

References

[1] S. Gerke, T. Minssen, and G. Cohen, "Chapter 12 - ethical and legal challenges of artificial intelligence-driven healthcare," in *Artificial Intelligence in Healthcare* (A. Bohr and K. Memarzadeh, eds.), pp. 295–336, Academic Press, 2020.

[2] P. Solanki, J. Grundy, and W. Hussain, "Operationalising ethics in artificial intelligence for healthcare: a framework for ai developers," *AI and Ethics*, 7 2022.

[3] A. Birhane, "Algorithmic injustice: a relational ethics approach," *Patterns*, vol. 2, p. 100205, 2 2021.

[4] K. Beck, J. Grenning, R. C. Martin, M. Beedle, J. Highsmith, S. Mellor, A. van Bennekum, A. Hunt, K. Schwaber, A. Cockburn, and et al., "Principles behind the agile manifesto," 2001.