Generative AI in Software Engineering Must Be Human Centered:

The Copenhagen Manifesto

PREAMBLE

The advent of Generative Artificial Intelligence—systems that can produce human-like content such as text, music, visual art, or source code—marks not only a significant leap for Artificial Intelligence (AI) but also a pivotal moment for software practitioners and researchers.

The role of software engineering researchers and practitioners in adopting the technologies that shape our world is critical. Historically, the human aspects of developing software have been treated as secondary to more technical innovations. However, the emergence of Generative AI will simultaneously enhance human capabilities while surfacing complex ethical, social, legal, and technical challenges.

While primarily aimed at software engineering (SE) researchers and practitioners, who are at the forefront of integrating Generative AI into our digital infrastructure, this manifesto equally underscores the ripple effects such innovations have on end-users of software, shaping their experiences and interactions in profound ways.

SE practitioners and researchers have a moral duty to help the world navigate this enduring landscape of new AI technologies ethically and sustainably, and foster values such as fairness, transparency, societal wellbeing, and environmental resilience.

This transformative era challenges us to extend our focus beyond technical expertise, integrating human values and ethical considerations into the fabric of our technological advancements This is in addition to maintaining our existing obligations, such as adherence to codes of conduct (e.g., ACM Code of Ethics and Professional Conduct[1] or the IEEE Code of Ethics[2]) and compliance with legal frameworks (e.g., EU AI Act[3]).

By establishing this manifesto, we seek to catalyze a shift in how Generative AI is conceived, developed, and applied within SE—a shift that reaffirms the primacy of human dignity, agency, and collective wellbeing in the face of rapid technological change.

1. https://www.acm.org/code-of-ethics
2. https://www.ieee.org/about/corporate/governance/p7-8.html
3. https://data.consilium.europa.eu/doc/document/ST-5662-2024-INIT/en/pdf

CORE VALUES OF THE COPENHAGEN MANIFESTO

1. Responsibility and Ethics: Emphasizing the duty to develop Generative AI in SE in a manner that does not bring (physical, emotional, or financial) harm to living beings.

2. Human-Centricity and Agency: Prioritizing human needs and autonomy when designing and using Generative AI in SE.

3. Transparency and Equity: Advocating for transparent, understandable, reproducible, and verifiable Generative AI in SE, including communications about research in this area, and equitable access and impact.

4. Inclusivity and Continuous Learning: Adopting, adapting, and empirically researching Generative AI in SE must inherently foster a culture of diversity, inclusivity, and continuous learning, embracing diverse, multidisciplinary perspectives for software engineering's advancement.

5. Environmental Sustainability: Committing to the responsible selection and use of Generative AI technologies, with a focus on minimizing environmental impact through energy efficiency and sustainable practices.

Bridging our core values to actionable principles, the Copenhagen Manifesto presents a roadmap for embedding these commitments into SE practices with Generative AI. These principles serve as concrete steps towards realizing a human-centered approach, ensuring that technological advancement is not only innovative but also equitable, transparent, and socially and environmentally responsible. As we move forward, these guidelines will steer us in applying Generative AI within SE in a manner that enhances the human experience, fostering an environment where technology truly serves society.

PRINCIPLES

1. Responsible Management of Generative AI in SE: Actively engage in responsibly managing and continuously evaluating Generative AI technologies, ensuring their alignment with ethical standards and societal needs.

2. Human Sovereignty Over AI: Ensure that Generative AI technologies enhance, rather than replace, human decision-making and creativity, prioritizing human oversight in Generative AI development and application.

3. Two Sides of Generative AI: Assess both the benefits and harms of Generative AI, promoting its use with caution and responsibility to avoid unintended consequences.

4. Sociotechnical Responsibility: Integrate social and technical considerations in the development of Generative AI-powered applications, aiming for solutions that are beneficial and respectful to all stakeholders.

5. Transparency and Fairness: Implement transparent processes that actively identify and mitigate Generative AI-related biases, ensuring fairness, accountability, and trustworthiness in Generative AI applications.

6. Sustainability and Environmental Impact: Select and advocate for Generative AI models and practices known for their lower environmental impact, emphasizing long-term sustainability.

PRINCIPLES

7. Equitable Impact of Generative AI in Software: Champion Generative AI that fosters inclusivity and equity, working to eliminate disparities and support empowerment across diverse communities.

8. Ethical Engineering and Education: Uphold ethical standards in both engineering practices and education, ensuring future generations of engineers are prepared to make ethical decisions in Generative AI use.

9. Complex Adaptive Nature of Generative AI: Recognize and plan the evolving nature of Generative AI systems by committing to ongoing learning and adaptation in their design and application, ensuring they remain aligned with ethical and societal values.

10. Empirical Evaluation of Professional Practices: Support and conduct empirical research to base Generative AI adoption and practices on evidence, enhancing the positive impact of Generative AI on society and addressing any challenges.

11. Enhancement and Wellbeing through Adaptive Generative AI in SE: Develop Generative AI that amplifies the capabilities of software engineers, promoting professional growth and wellbeing.

12. Public Awareness and Discourse: Engage in and foster public discussions about Generative AI, aiming to educate and inform while combating misinformation, thereby shaping a balanced and factual narrative around Generative AI in SE.

CONCLUSION

As we explore the possibilities of Generative AI in SE, we must keep our eyes open not only to its incredible opportunities but also to its considerable risks.

Generative AI is reshaping the landscape, affecting how we live and work. Our mission is to steer Generative AI in SE towards fairness, transparency, and justice, ensuring a future where technology lifts us up together, rather than divides us.

Talking about responsibility, ethics, transparency, fairness, and focusing on everyone's wellbeing means insisting that all technology, including Generative AI, serves the common good.

It is about developing and applying Generative AI with respect for each person's rights and dignity to build a world that is better for our collective effort.

In conclusion, this manifesto represents our collective commitment to a future where technology is guided by the values and principles we have laid out.

Also, it is a pledge to advance software engineering not just as a field of technical expertise but as a domain profoundly connected to ethical integrity, social responsibility, and human wellbeing.

Acknowledgments

Our sincere appreciation goes to the Alfred P. Sloan Foundation and the Carlsberg Foundation for their invaluable support of the Copenhagen Symposium on Human-Centered Software Engineering AI (Grants G-2023-21020 and CF23-0208 to Daniel Russo). Held at Aalborg University – Copenhagen from November 26-29, 2023, this symposium was the birthplace of our Manifesto, and contributions from these Foundations have been crucial in shaping its vision and direction.

Participants

1. Daniel Russo - Aalborg University, Denmark
2. Sebastian Baltes - University of Bayreuth, Germany
3. Niels van Berkel - Aalborg University, Denmark
4. Paris Avgeriou - University of Groningen, The Netherlands
5. Fabio Calefato - University of Bari, Italy
6. Beatriz Cabrero-Daniel - University of Gothenburg & Chalmers, Sweden
7. Gemma Catolino - University of Salerno, Italy
8. Jürgen Cito - TU Wien, Austria
9. Neil Ernst - University of Victoria, Canada
10. Thomas Fritz - University of Zurich, Switzerland
11. Hideaki Hata - Shinshu University, Japan
12. Reid Holmes - University of British Columbia, Canada
13. Maliheh Izadi - Delft University of Technology, The Netherlands
14. Foutse Khomh - Polytechnique Montreal, Canada
15. Mikkel Baun Kjærgaard - University of Southern Denmark, Denmark
16. Grischa Liebel - Reykjavik University, Iceland
17. Alberto Lluch Lafuente - Technical University of Denmark, Denmark
18. Stefano Lambiase - University of Salerno, Italy
19. Walid Maalej - University of Hamburg, Germany
20. Gail Murphy - University of British Columbia, Canada
21. Nils Brede Moe - SINTEF, Norway
22. Gabrielle O'Brien - University of Michigan, USA
23. Elda Paja - IT University of Copenhagen, Denmark
24. Mauro Pezzè - USI Università della Svizzera Italiana, Switzerland
25. John Stouby Persson - Aalborg University, Denmark
26. Rafael Prikladnicki - Pontifícia Universidade Católica do Rio Grande do Sul, Brazil
27. Paul Ralph - Dalhousie University, Canada
28. Martin Robillard - McGill University, Canada
29. Thiago Rocha Silva - University of Southern Denmark, Denmark
30. Klaas-Jan Stol - University College Cork, Ireland
31. Margaret-Anne Storey - University of Victoria, Canada
32. Viktoria Stray - University of Oslo, Norway
33. Paolo Tell - IT University of Copenhagen, Denmark
34. Christoph Treude - Singapore Management University, Singapore
35. Bogdan Vasilescu - Carnegie Mellon University, USA