Software Can Do Wrong: On Ethics in Agile Software Engineering

Alexander Pretschner Software&Systems Engineering@TUM – bidt – fortiss – CDTM

jww Jan Gogoll, Severin Kacianka, Julian Nida-Rümelin, Niina Zuber DigHum Lectures – May 16th, 2023



Bavarian Research Institute for Digital Transformation



Mission: Understand and help design digital transformation in society

Interdisciplinarity, networking, dialogic research

Three departments: research, think tank, dialog

Working on: digital competition, ethics&law in digital transformation, power of opinion, digital economies, digital work environments

Established 2019, now ~32 people

www.bidt.digital





Context:

Digital Humanism – how "responsible" can a machine be? I—a software engineer—believe in personal responsibility.

Don't offload responsibility to others or "the system". Yet, innovation and money matter, too.

"Ethics" is not an AI concern only. It's a *software* concern. Regulation later today.

So what *could* we do?

- Education: Raise awareness. Tech students love it. Effects unclear.
- Research: Impact statements when submitting papers, IRBs or ERBs, ...
- Certification: For (self-educated) software engineers? For companies?
- Development: IRBs. Ethical Deliberation in Agile Development

• ...





Diffusion of responsibility: from research to practice



[Taken from: Hanna Wallach, keynote at NeurIPS 2020, https://nbiair.com/#Recordings]

EU regulation: (high risk) AI applications



17. Mai 2023

4

Ethics in Software Engineering

Reproach to ethicists: "Useless!" (and to software engineers: "Not informatics anymore!")

Indeed: >120 Codes of Conduct for AI/Software/Systems Engineering rather fruitless

Reason: Software context-specific; hence values and trade-offs context-specific w.r.t. application domain, technology, users' culture, developers' culture, optimization goals, ...

Examples: face recognition, data integration, care robots, resume analyzers, etc. – but also software without AI: ego shooters, Corona app, BitTorrent, Telegram, Bitcoin, website preferences, ...

Genericity of CoCs hence necessary. Only way out: Deliberation schema that caters to context specificity.



Plus: What about trade-offs?



Gogoll, Zuber, Kacianka, Greger, Pretschner, Nida-Rümelin: Ethics in the Software Development Process: from Codes of Conduct to Ethical Deliberation. *Philos. Technol.* 34: 1085–1108, 2021 <u>https://link.springer.com/article/10.1007/s13347-021-00451-w</u>

CoC necessarily generic - McNamara et al. (2018) find no evidence that CoCs influence behaviour

Software Engineering and AI

Ethical issues are not confined to AI – but this is suggested by the current debate! A centralized Corona app? Infrastructure like Palantir Foundry? Integration of registers?



Sculley et al.: Hidden Technical Debt in ML systems, Proc. NIPS 2015: 2503-2511



Ethical Deliberation in Agile Processes

https://www.bidt.digital/wp-content/uploads/2021/04/Digital-Transformation-and-Ethics_Zuber-et-al_EN.pdf

No simple way out. Need to address concerns in a context-specific manner: think! Can be done in a systematic way

Development driven by EDAP: "Ethical Deliberation in Agile [Development] Processes"

Key idea: start with and iterate on values; continuously reflect on **mechanisms to implement them, not yes/no**

Characteristics of agility blend particularly well: planning; incrementality; empowerment; learning

Zuber, N., *et al.* Empowered and embedded: ethics and agile processes. *Humanit Soc Sci Commun* **9**, 191 (2022) <u>https://www.nature.com/articles/s41599-022-01206-4</u>



Scrum





9

Scrum in Context



Relevant categories of agility

The Role of Planning: Undoing the Separation of Design and Production

Empowerment

Incrementality and 100%-artifacts

Retrospectives and learning

(and sure there's more: easiness, agile manifesto)



Agility: Power to the People!



bidt Ein Institut der Bayerischen Akademie der Wissenschaften

Systematic Techno-Ethical Deliberation (EDAP)

Activity I: Descriptive System Analysis

Activity II: (Descriptive) Value Analysis: Company, Professional, Techno-Generic, Data-Generic, Domain-Specific, Context-Specific Values

Activity III: Techno-ethical Deliberation: From Values to Requirements to Mechanisms

Activity IV: Ethical System Review

Activity V: Verification

Zuber, N., Kacianka, S., Pretschner, A. & Nida-Rümelin, J. (2020). *Ethische Deliberation für agile Softwareprozesse: EDAP-Schema* (Band Digitale Transformation und Ethik). Ecowin.

bickt Ein Institut der Bayerischen Akademie der Wissenschafter

Ethical Deliberation in Scrum





Examples: (Un)Ethical Pizza Delivery Apps and ChatGPT Dolls







15

Who is responsible?



Certification?



Individual

Process

Organization



Final Comments

Software is unique: invisibility, malleability, scalability, many-hands, networked

Teaching ethics as part of software development classes

Ethics not reducible to compliance: there the default is to see people as risk! Empowerment?

Power asymmetry: programmers can quit

Considerations apply to both UI and program logic

ML as one "virtual" sprint

DevOps, specifically MLOps: ethical consideration doesn't stop at a specific moment in time. For AI, context continuously changes

Is ethics a first order problem?



Example: ChatGPT Doll



ChatGPT Doll

Next Generation Doll

Connected to the Internet

Responds to Voice Commands

Engages in Basic Conversation



