Embedded Ethics as Proactive Regulation of Technology

(A New Solution to the Collingridge Dilemma?)

Daniel W. Tigard, PhD

Institute for History and Ethics of Medicine
Technical University of Munich
daniel.tigard@tum.de

bidt Ein Institut der Bayerischen Akademie der Wissenschaften

ПUП

Key Questions

- What does it mean to "embed" ethics into technology?

- How, if at all, can embedded ethics provide proactive (or preparatory) regulation?

- Might this help to address the "Collingridge Dilemma"?

Big Data, Artificial Intelligence, & "Novel Beings"

"Embedded Ethics" (EE)

- "The practice of integrating the consideration of social, ethical and legal issues into the entire development process in a deeply integrated, collaborative and interdisciplinary way."

nature machine intelligence

Search 🔍 Login ⊗

Comment | Published: 31 July 2020
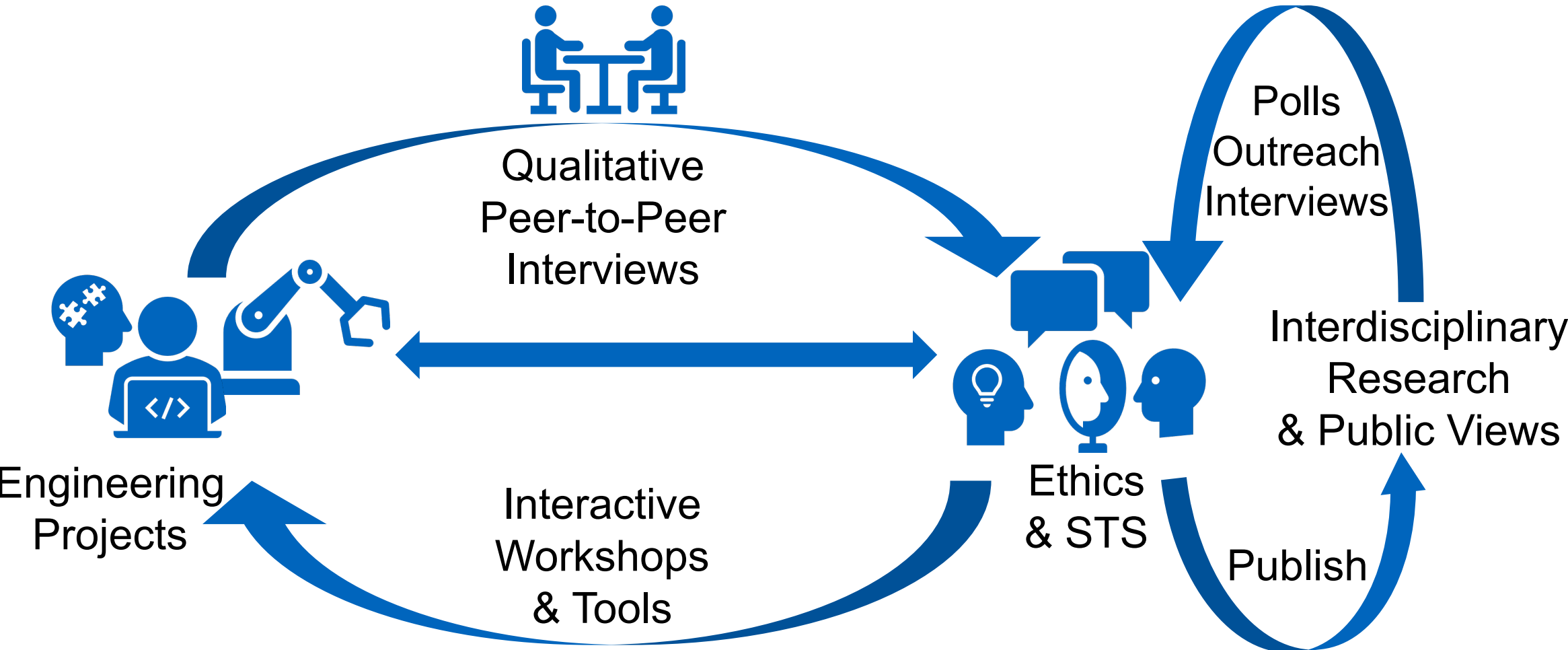
An embedded ethics approach for AI development

Stuart McLennan ✉, Amelia Fiske, Leo Anthony Celi, Ruth Müller, Jan Harder, Konstantin Ritt, Sami Haddadin & Alena Buyx

Nature Machine Intelligence (2020) | Cite this article

86 Accesses | **74** Altmetric | Metrics

bidt Ein Institut der Bayerischen Akademie der Wissenschaften

TUM

4

EE: Structure + Methods
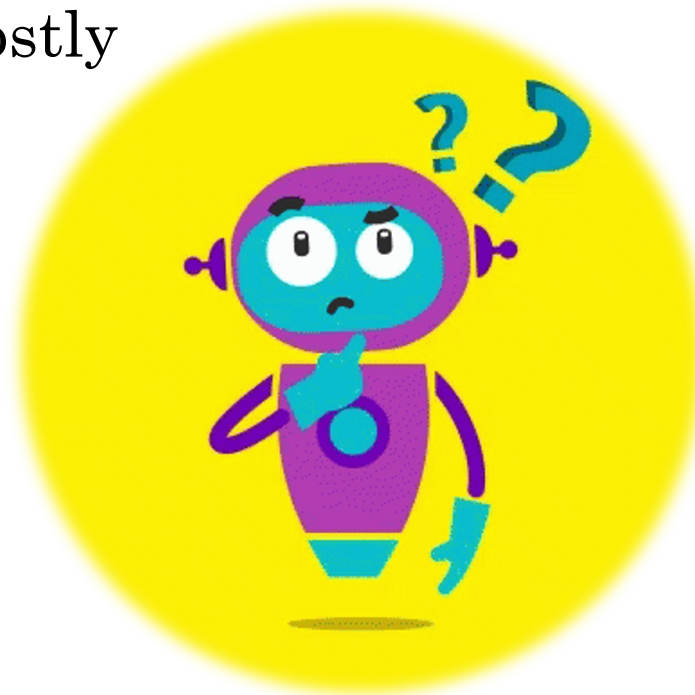
Proactive/Preparatory Regulation

- The attempt to mitigate in advance the potentially negative impacts of technology, and to secure the potentially positive impacts, based upon prediction or ongoing awareness of technology's emergence

"Collingridge Dilemma"

- "during its early stages, when it can be controlled, not enough can be known about its harmful social consequences to warrant controlling its development; but by the time these consequences are apparent, control has become costly and slow."
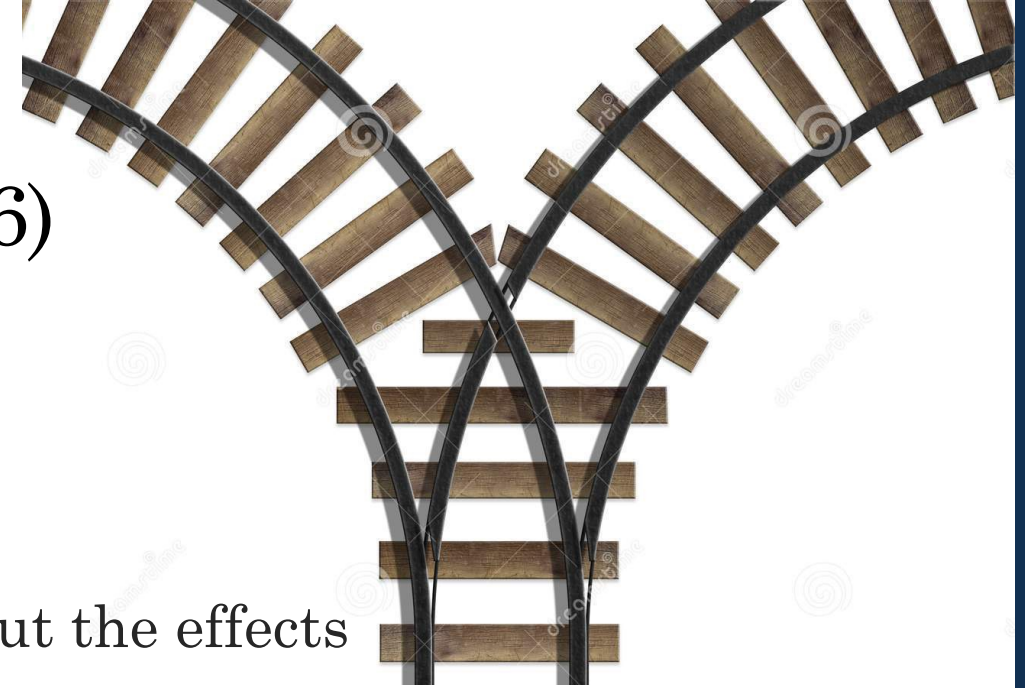 (Collingridge 1980)

2 Interpretations (Croy 1996)

Regulation is either…

- Impossible
  - b/c (now) we don't know enough about the effects
  - (later) it's too costly/slow to help

- Doomed to fail
  - b/c (now) we can't reliably predict
  - (later) tech is "societally embedded" (Kudina and Verbeek 2019)

"Embedded Ethics" as Proactive/Preparatory

- Embedding Ethics ≠ Embedding Ethic<u>ists</u>

- Goal = foster sincere *sensitivity* in developers themselves, an ability to recognize and effectively think through the social & ethical features of the systems under development.

- Our role is exploratory and assistive
  - Not *monitoring* or *assessing*

- Our role is perhaps temporary

Sensitivity as Flexibility

- Flexibility in developers: remaining open to new future options (Collingridge)
  - But again, problem of predictability

- Flexibility: an ability to adapt to new circumstances **and** newly discovered perspectives and values brought about by our use of technology

Regulating without Regulation

- EE = building competencies in developers for identifying and addressing non-technical features of their work.

- These goals are best met by working collaboratively, rather than monitoring or assessing.

- By avoiding explicit attempts to regulate, we can contribute to the effort to mitigate potentially negative impacts and to secure potentially positive impacts.

- Thus, by avoiding explicit regulation, we can implement preparatory modes of regulation.

bidt Ein Institut der Bayerischen Akademie der Wissenschaften

Thank you!

Daniel W. Tigard, PhD

Institute for History and Ethics of Medicine
Technical University of Munich
daniel.tigard@tum.de