Project title: Methods of Ethics for Socially Disruptive Technologies

May 2021

Type of project: 3-year postdoc

Intended starting date: September 1st, 2021

Name daily supervisor(s), affiliation, and intended promotor in case of PhD student:
Joel Anderson and Sven Nyholm (Utrecht University)

Involved ESDT-partner(s) from other universities (if any):
Philip Brey (Twente), Wijnand Ijsselsteijn (TU/e, for coordination with the STEM-track postdoc and Human Line), Marcel Verweij (WUR, for coordination with the Nature line)

The Research Fellows and Postdocs of the Foundations & Synthesis research line, including Elizabeth O’Neill (TU/e), Julia Hermann (Twente), Michael Klenk (Delft), Guido Löhr (TU/e) and Jeroen Hopster (Twente).

Proposed (and approved) Research Project:
The aim of this postdoc project is to investigate the following research question:

“Which ethical methods are best suited (and on what grounds) for understanding and adjudicating the normative concerns raised by emerging technologies (particularly, their socially and conceptually disruptive character)?”

This project addresses a complex and multi-faceted range of issues.

- Some socio-technological developments may require fundamental rethinking of our normative accounts, while others can be addressed using familiar frameworks.
  - Methods for analyzing ethical issues and justifying ethical judgments been around for a long time, but only recently have specific methods been proposed for emerging technologies.

- Emerging and socially disruptive technologies may share distinctive ethical characteristics, but they may also differ fundamentally in the methods needed to analyze them. These technologies raise some distinct methodological issues, particularly in virtue of their dynamic uncertainty:
  - uncertainty about what social disruptions they will generate,
  - uncertainty about new and self-replicating forms of agency they produce; and
  - uncertainty about how the disruption of our ethical concepts may also transform our normative perspectives themselves.

- And the discussion of these issues occurs against a background of ongoing disputes about methodological pluralism and the relationship between more
contextual methods and more principled methods in ethics, political philosophy, and related fields of philosophical and scientific inquiry.

- There is growing commitment in ethics (and, more broadly, “practical philosophy”) to both attending to the specificity of the contexts in which technologies are employed and also recognizing the diversity of perspectives that can be relevant for ethical judgments.
- At the same time, as the disruptive effects of technologies increasingly disregard the boundaries of individual countries, the need for shared context-transcending principles of ethical governance are also becoming more urgent.
- Assessing the advantages and disadvantages of methodological pluralism is further complicated by differences between approaches withings ethics and practical philosophy as well, which includes not only normative ethics, metaethics, and applied ethics, but also political theory, human rights law, philosophical anthropology, critical social science, feminism, technology studies, postphenomenology, and moral psychology.

This postdoc project will focus on clarifying what is most suitable, in terms of normative justification and methods of ethics and practical philosophy, for meeting the distinctive challenges of evaluating socially disruptive emerging technologies. In alignment with one of the central aims of the “Ethics of Socially Disruptive Technology” consortium, this project will be open to a diverse range of methodologies, in terms of what could be termed a “methodologically pluralistic practical philosophy of technology” that – as an ideal – aspires to be informed by:

- the full range of sub-disciplinary approaches within practical philosophy;
- a close involvement with the relevant empirical work, especially in the behavioral and social sciences, evolutionary biology, artificial intelligence, neuroscience, and human-machine interface;
- a deep engagement with perspectives at the margins (or outside) of mainstream Western philosophy, including “comparative philosophy” (non-Western philosophical traditions), feminism, critical race theory, disability studies, and a variety of “situated epistemologies” that shed new, normatively relevant light on issues of socially disruptive technologies.

A key part of this task is to develop what one might think of as the “design specifications” for the normative methodologies that can provide guidance in various projects undertaken by the consortium. Which approach to normative justification and analysis, in other words, is needed for meeting the evaluative and diagnostic challenges that are specific to socially disruptive emerging technologies?

What follows is a preliminary indication of the constellation of relevant possible questions – from which a suitable subset would be selected for the actual project, depending on the interests and specialization of the postdoc.

- To what extent (if at all) is there a common denominator in ethical questions raised by socially disruptive technologies?
- To what extent should normative principles incorporate contextual factors regarding the technologies being developed (such as how foreseeable their development is, how socially vulnerable the intended users are, what the cultural context of application is, or what the involvement of stakeholders is)?
- In determining what forms of “social disruption” needs evaluation, is the target of evaluate best seen in terms of a matter of states of affairs (“impact”), or in terms of violation of norms, implications for human capabilities or relationships?
- Does an integrative approach necessarily have to have a unitary ethical approach (say, consequentialism), or is methodological pluralism appropriate (in which approaches are combined and contextualized)?
• What is the relationship between more abstract and principled “moral point of view” and the codification of principles in law, policies, social impact assessment procedures, professional codes, etc.?
• To what extent ought the relevant normative methods permit formalization or operationalization, either in protocols or decision-supporting algorithms?
• In what ways ought normative approaches to socially disruptive technologies be responsive to empirical findings, for example regarding what is politically feasible, psychologically realistic, technologically possible, or socially accepted?

In parallel with this meta-normative research on methods, there is also a need for reflecting on these methodological issues in collaboration with projects being carried out by members of the consortium. The postdoc would be expected to collaborate with at least two ESDT PhD projects, which can serve as contexts for in-depth reflection on methodological issues, in projects in the Nature, Society, and Human research lines. These could either be as a testbed for applying the methodological recommendations developed elsewhere and/or as a methodological “red team,” serving as a designated methodological gadfly. This embedded reflective perspective would be particularly relevant in prioritizing the methodological issues that are especially relevant, such as examining which normative methods are needed for assessing disruptive impact of algorithms that assist with ethical decision-making.

Relation to ESDT research lines and sublines:

At a most fundamental level, the “Ethics of Socially Disruptive Technology” is concerned with the normative evaluation of the (anticipated) effects of technologies.

In particular, this postdoc will contribute to the research agenda items (2) “Ethical analysis of SDTs” and (4) “Reconceptualizations and the research agenda of practical philosophy and other fields” as found in the gravitation grant proposal. The postdoc will strengthen the Synthesis & Foundations line by working closely with the postdocs on concept change and socially disruptive technologies, by highlighting the issues of normative justification central to determining which conceptual changes are for the better and whether a social disruption is normatively problematic or a welcome transformation.