



HELMUT SCHMIDT  
UNIVERSITÄT

Universität der Bundeswehr Hamburg

# ETH TECH



**Co-funded by  
the European Union**

## What is Tech/AI ethics for us?

Sigrid Hartong &  
Ina Sander  
(HSU Hamburg)

KOM ETH-TECH  
Padua, 22-24 Jan '25

# Ethics

Ethics = values and norms that affect our everyday decision making or that we want to affect our everyday decision making (problem of, e.g., institutional norms vs. personal norms)

Ethical view on technology = matching what tech “is” and “does” (= what is triggered through tech) with the personal/institutional values and norms, and aligning tech implementation/usage to the personal/institutional norms

# Observation we share



- Even though there is a dominance of techno-optimism, many actors mention critical aspects of AI and hereby use the term „ethics“ very often
- However, „indepth reflections are rare. This approach can serve as a rhetorical strategy to pre-empt potential criticism without fully addressing the underlying structural issues that have contributed to the very problems technologies are supposed to solve. Furthermore, the narratives often imply the inevitability of innovation through digitalisation, diverting attention away from possible alternatives to technological solutions.”

(Shi/Palenski 2024: 5)



# Our approach to ethics: tech/AI „ethics“ need to adress/be grounded in (knowing about) broader interrelations and planetary structures/effects

## Technology, Activism and Living among Planetary Ruins

Felicitas Macgilchrist

### 1 Introduction

Transformation holds a deep fascination for many people: “Digital transformation” and “ecological transformation” are major themes in current public, policy and economic discourse. In education, interventions often flow towards what young people should *know*, how they should change their *attitudes*, what they should *do*. When digital technologies are mentioned in research in this context, they are often understood primarily as “tools” to support changes in understanding, attitudes, beliefs, values and lifestyle behaviour (Ahel & Lingenau, 2020; Al-Mulla, Ari, & Koç, 2022; Hallinger et al., 2020; Janakiraman et al., 2021; Spangenberg et al., 2020). Software and websites abound to support young people in critically engaging with their own consumer behaviour, or encouraging them to save heat, water and electricity. Gamified apps invite them to calculate their carbon footprint, set goals through actions such as buying second-hand rather than new, driving less, talking more about climate or celebrating local ‘Earth Heroes’.<sup>1</sup>

But are these kinds of change sufficient for the multiple crises we are currently facing? In this chapter, aligning with an emerging body of critical work on the nexus of educational technology (edtech) and degrowth, low tech or other systemic transformations, my response is: no (see, e.g., Grünberger, 2021; Grünberger et al., 2021; Macgilchrist, 2021; Scherrer, 2022; Selwyn, 2021; Selwyn et al., 2020). The kind of research and edtech noted in the previous paragraph, which arguably takes an “instrumental approach” to ESD (Getzin & Singer-Brodowski, 2016), does not address the planetary dimensions. Precisely because of this, they promote what Ingolfur Blühdorn and colleagues (2020) call a “sustainable non-sustainability”. I will suggest

<sup>1</sup> Materials or apps are available, for instance, here: <https://www.earthhero.org>, <https://klima.com>, <https://www.regenwald-schuetzen.org/projekte/bildungs-projekte/abgeschlossene-projekte/projekte-fuer-die-sekundarstufe-1/erdkunde-biologie/systeme-verstehen/mitmachkrimi-tatort-tropenwald-20> (Last access: 22.11.2023)

- ◆ Importance of moving towards a “planetary view” on (ed)tech (Macgilchrist 2024):
  - ◆ Ecological dimensions
  - ◆ Global capitalist dimensions (production chains, data capitalism,...) (see also Knox 2019)
  - ◆ (Geo)Political dimensions
  - ◆ Social Dimensions (reproduction of inequality, privilege hazards,...)
- ◆ Values and norms, then, would include: how to live on planetary “ruins”, peace, caring (better) for the vulnerable, fairer work, community orientation, addressing conflict,...
- ◆ Importance of questioning (ed)tech with regards to these values, but also with regards to education/learning that is (not) being triggered through tech

1



## Human Agency and Oversight

- Is the teacher role clearly defined so as to ensure that there is a teacher in the loop while the AI system is being used? How does the AI system affect the didactical role of the teacher?
- Are the decisions that impact students conducted with teacher agency and is the teacher able to notice anomalies or possible discrimination?
- Are procedures in place for teachers to monitor and intervene, for example in situations where empathy is required when dealing with learners or parents?
- Is there a mechanism for learners to opt-out if concerns have not been adequately addressed?
- Are there monitoring systems in place to prevent overconfidence in or overreliance on the AI system?
- Do teachers and school leaders have all the training and information needed to effectively use the system and ensure it is safe and does not cause harms or violate rights of students?

2



## Transparency

- Are teachers and school leaders aware of the AI methods and functions?
- Is it clear what aspects AI can take over and what not within the classroom?
- Do teachers and school leaders understand how specific aspects of the AI system work?
- Are the system processes and outcomes focussed on the expected predictions, assessments and classifications of the AI system?

3



## Diversity, non-Discrimination and Fairness

- Is the system accessible by everyone in the same way without any barriers?
- Does the system provide appropriate interaction modes for learners with disabilities or special education needs? Is the AI system designed to treat learners respectfully adapting to their individual needs?
- Is the user interface appropriate and accessible for the age level of the learners? Has the usability and user-experience been tested for the target age group?
- Are there procedures in place to ensure that AI use will not lead to discrimination or unfair behaviour for all users?
- Does the AI system documentation or its training process provide insight into potential bias in the data?
- Are procedures in place to detect and deal with bias or perceived inequalities that may arise?

4



## Societal and Environmental Wellbeing

- How does the AI system affect the social and emotional wellbeing of learners and teachers?
- Does the AI system clearly signal that its social interaction is simulated and that it has no capacities of feeling or empathy?
- Are students or their parents involved in the decision to use the AI system and support it?
- Is data used to support teachers and school leaders to evaluate student wellbeing and if so, how is this being monitored?
- Does use of the system create any harm or fear for individuals or for society?

5



## Privacy and Data Governance

- Are there mechanisms to ensure that sensitive data is kept anonymous? Are there procedures in place to limit access to the data only to those who need it?
- Is access to learner data protected and stored in a secure location and used only for the purposes for which the data was collected?
- Is there a mechanism to allow teachers and school leaders to flag issues related to privacy or data protection?
- Are learners and teachers informed about what happens with their data, how it is used and for what purposes?
- Is it possible to customise the privacy and data settings?
- Does the AI system comply with General Data Protection Regulation?

6



## Technical Robustness and Safety

- Is there sufficient security in place to protect against data breaches?
- Is there a strategy to monitor and test if the AI system is meeting the goals, purposes and intended applications?
- Are the appropriate oversight mechanisms in place for data collection, storage, processing, minimisation and use?
- Is information available to assure learners and parents of the system's technical robustness and safety?

7



## Accountability

- Who is responsible for the ongoing monitoring of results produced by the AI system and how the results are being used to enhance teaching, learning and assessment?
- How is the effectiveness and impact of the AI system being evaluated and how does this evaluation consider key values of education?
- Who is responsible and accountable for final decisions made regarding the procurement and implementation of the AI system?
- Is there a Service Level Agreement in place, clearly outlining the support and maintenance services and steps to be taken to address reported problems?

<https://education.ec.europa.eu/news/ethical-guidelines-on-the-use-of-artificial-intelligence-and-data-in-teaching-and-learning-for-educators>



...the key message: there is no one best scenario,  
because there are always dimensions that are  
„suffering“

Ecological dimensions

Global capitalist dimensions  
(production chains, data capitalism,...)

(Geo)Political dimensions

Social Dimensions (reproduction of  
inequality, privilege hazards,...)

Learning Dimensions

Example: Building „own“ AI  
servers/LLMs to create alternatives to  
OpenAI  
→ more infrastructure that requires  
servers, energy for LLM training until  
it is „as good“ as ChatGPT,...

Example: AI tool „helps“ students to  
produce creative output (e.g., images)  
but they are simultaneously „drawn  
into“ the product and depend on  
global capitalism and the power of  
few



Knowing this can result in  
disenchantment, but it does not  
have to if it is viewed from a “small  
steps” perspective and driven by a  
strong vision of the future

# An approach to bridge everyday practices, ethical values and tech/AI → „An ethics of practice“ (Bezuidenhout et al. 2020: 2197)

- Making “visions of the future” more strongly visible (→ AWS), including their ethical dimensions (which kind of society/university/learning of the future do we want?)
- Making “ethics of practice” more strongly visible → Bringing visions/values in relation to a planetary view of tech/AI as well as in relation to everyday (university) practices (e.g., *creating 2 images with AI means boiling 3 water kettles...; shifting teaching to MS Teams means more power centralization on big tech and dependency on MS design of organizing classes*)
- Developing ideas for alternative practices (“microethics”, Bezuidenhout/ Ratti 2021: 939) that do not overcome dilemmas, but that (better) contribute to the visions of the future

Science and Engineering Ethics (2020) 26:2189–2213  
<https://doi.org/10.1007/s11948-020-00197-2>

ORIGINAL RESEARCH/SCHOLARSHIP



## “Ethics When You Least Expect It”: A Modular Approach to Short Course Data Ethics Instruction

Louise Bezuidenhout<sup>1</sup> · Robert Quick<sup>2</sup> · Hugh Shanahan<sup>3</sup>

Received: 13 February 2019 / Accepted: 10 February 2020 / Published online: 17 February 2020  
© The Author(s) 2020

### Abstract

Data science skills are rapidly becoming a necessity in modern science. In response to this need, institutions and organizations around the world are developing research data science curricula to teach the programming and computational skills that are needed to build and maintain data infrastructures and maximize the use of available data. To date, however, few of these courses have included an explicit ethics component, and developing such components can be challenging. This paper describes a novel approach to teaching data ethics on short courses developed for the CODATA-RDA Schools for Research Data Science. The ethics content of these schools is centred on the concept of open and responsible (data) science citizenship that draws on virtue ethics to promote ethics of practice. Despite having little formal teaching time, this concept of citizenship is made central to the course by distributing ethics content across technical modules. Ethics instruction consists of a wide range of techniques, including stand-alone lectures, group discussions and mini-exercises linked to technical modules. This multi-level approach enables students to develop an understanding both of “responsible and open (data) science citizenship”, and of how such responsibilities are implemented in daily research practices within their home environment. This approach successfully locates ethics within daily data science practice, and allows students to see how small actions build into larger ethical concerns. This emphasises that ethics are not something “removed from daily research” or the remit of data generators/end users, but rather are a vital concern for all data scientists.

✉ Louise Bezuidenhout  
[Louise.bezuidenhout@insis.ox.ac.uk](mailto:Louise.bezuidenhout@insis.ox.ac.uk)

Robert Quick  
[rquick@iu.edu](mailto:rquick@iu.edu)

Hugh Shanahan  
[Hugh.Shanahan@rhul.ac.uk](mailto:Hugh.Shanahan@rhul.ac.uk)

<sup>1</sup> Institute for Science, Innovation and Society, University of Oxford, Oxford, UK



# Principles

1. Moving away from “solving” things/checkbox thinking: “the way in which these terms [e.g., ethics] are widely coupled in literature rests on a (more or less unacknowledged) essentialist and axiomatic ground which ultimately oversimplifies the issue at stake – how can we do good with/through data-driven technologies? – and hinders an effective tackling of the moral dilemmas that data-driven technologies pose when adopted in complex real-life scenarios.” (Calzati/Ploeger 2024: 1)
2. Ethics not as “toolbox”, but as practice and, particularly, as “method of [ongoing] inquiry” (for which everybody is authorized and responsible) (ibid.: 2)
3. Ethics as non-axiomatic (it’s not either good or bad, rather e.g., bad for whom in which context?)

Original Research Article



## Problem-solving? No, problem-opening! A method to reframe and teach data ethics as a transdisciplinary endeavour

Stefano Calzati<sup>1</sup> and Hendrik Ploeger<sup>1</sup>

### Abstract

Starting from the recognition of the limits of today's common essentialist and axiological understandings of data and ethics, in this article we make the case for an ecosystemic understanding of data ethics (for the city) that accounts for the inherent value-laden entanglements and unintended (both positive and negative) consequences of the development, implementation, and use of data-driven technologies in real-life contexts. To operationalize our view, we conceived and taught a master course titled 'Ethics for the data-driven city' delivered within the Department of Urbanism at the Delft University of Technology. By endorsing a definition of data as a sociotechnical process, of ethics as a collective practice, and of the city as a complex system, the course enacts a transdisciplinary approach and problem-opening method that compel students to recognize and tackle the unavoidable multifacetedness of all ethical stances, as well as the intrinsic open-endedness of all tech solutions, thus seeking a fair balance for the whole data-driven urban environment. The article discusses the results of the teaching experience, which took the form of a research-and-design workshop, alongside the students' feedback and further pedagogical developments.

Big Data & Society  
July–September: 1–14  
© The Author(s) 2024  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/20539517241270687  
journals.sagepub.com/home/bds



Central “game changer”: from making people fear ethical mistakes (= desire to check boxes to feel better) towards wanting to inquire on ethics and bringing findings actively into the collective discussion (yet, must be mirrored by vision of a learning organization!!)



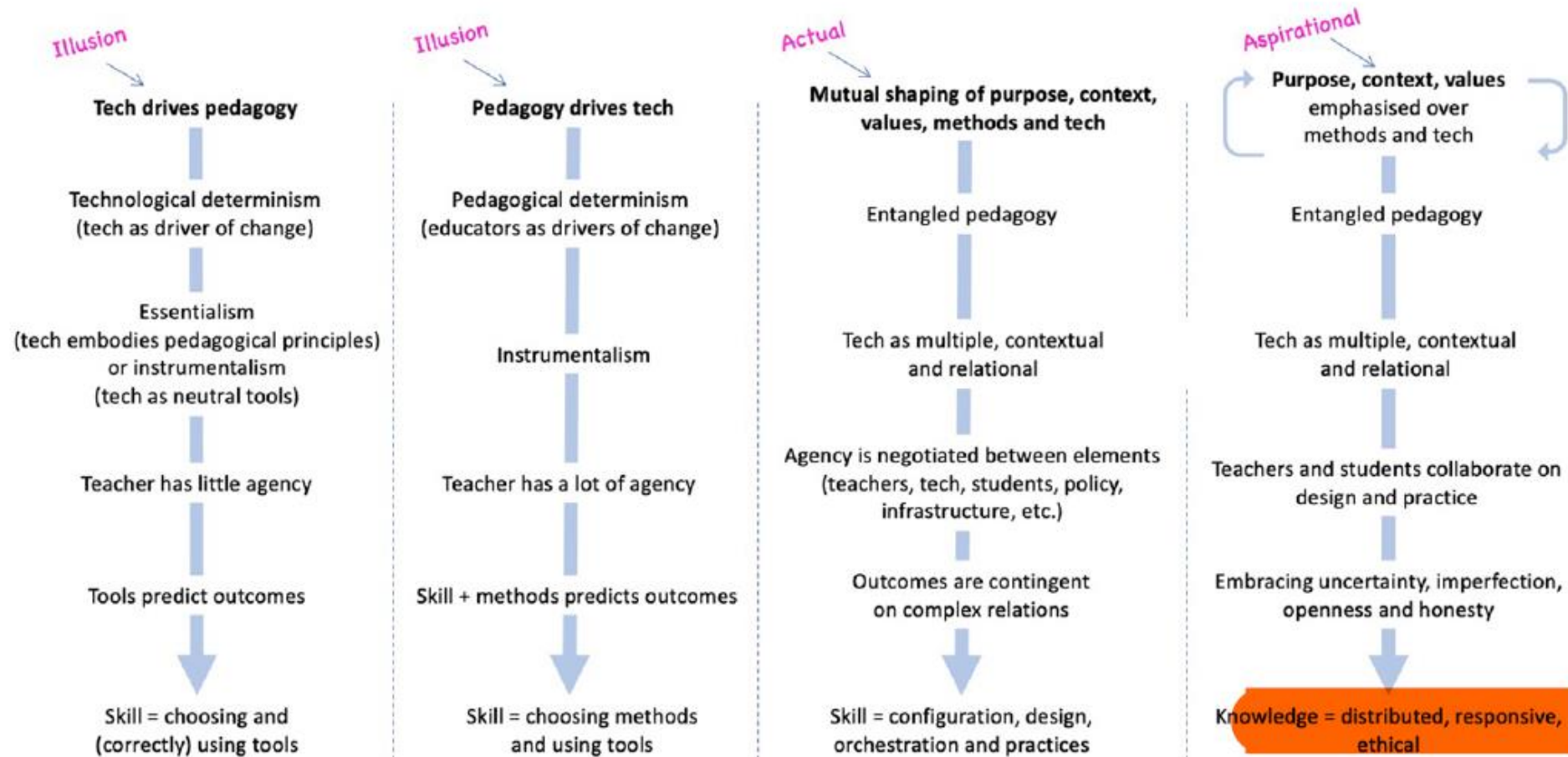
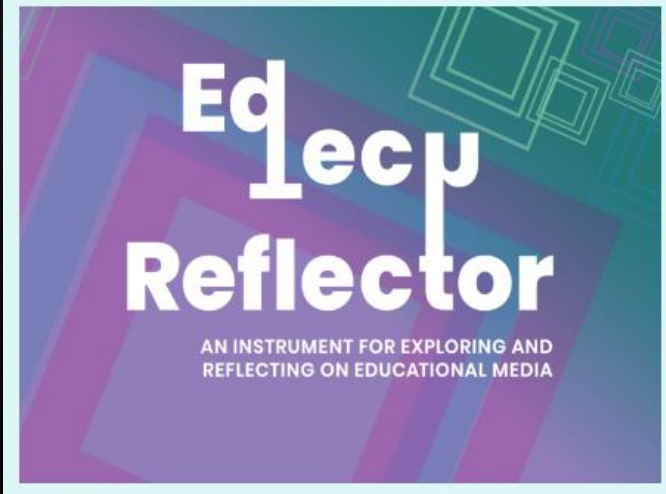
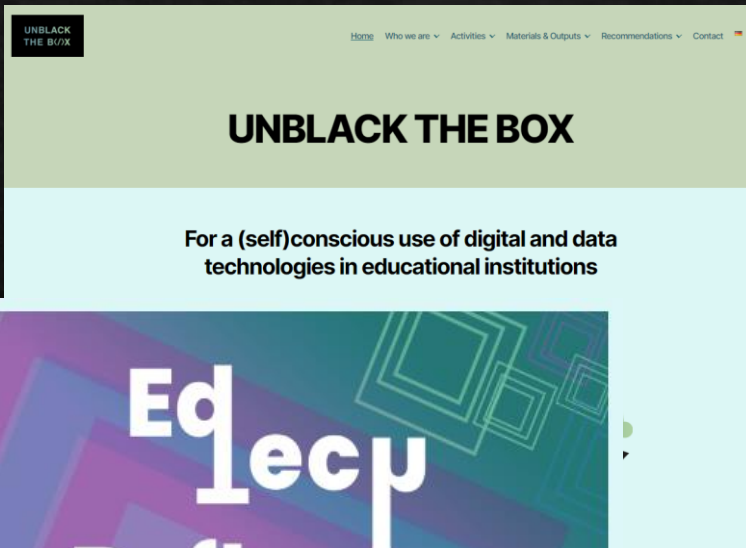


Fig. 2 An entangled pedagogy, including an aspirational view

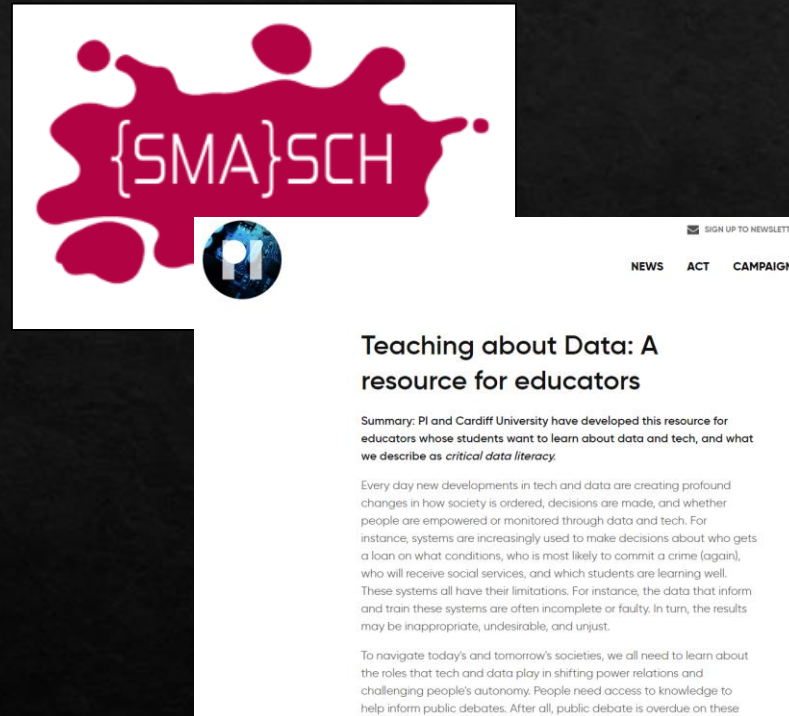
# How we so far try to support these activities in our transfer and co-design projects...

www.smasch.eu

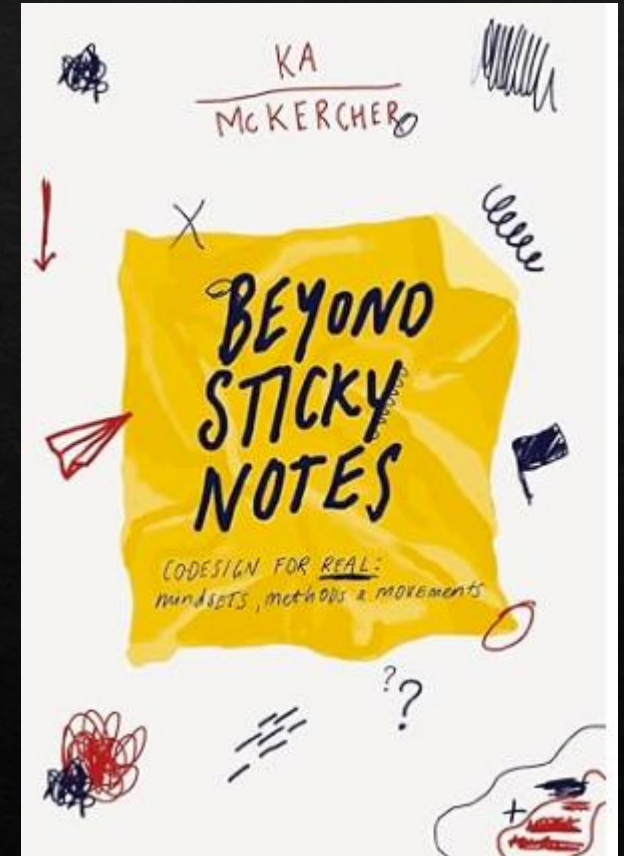


<https://unblackthebox.org>

<https://unblackthebox.org/materialien-ergebnisse/edtechreflektor/>



<https://privacyinternational.org/learning-resources/teaching-about-data-resource-educators>





# This would, then, also inform our ETH-TECH OER development

- Making “visions of the future” more strongly visible (which kind of society/university/learning of the future do we want?)
- Making “ethics of practice” more strongly visible
  - Bringing visions/values in relation to a planetary view of tech/AI as well as in relation to everyday practices
- Developing ideas for alternative practices (“microethics”)

- Openly addressing moral dilemmas of complex real-life scenarios and **not trying to find simple solutions**
- Ethics as practice and, particularly, as “**method of [ongoing] inquiry**” (**in which we are all together**)
- Ethics as **non-axiomatic** (it’s not either good or bad, rather e.g., bad for whom in which context?)

- Designing the tools itself can be regarded as practices of ethical inquiry, in which many people should, consequently, participate with different positions/ideas

# References

- ◆ Bezuidenhout, L., Quick, R., & Shanahan, H. (2020). “Ethics when you least expect it”: a modular approach to short course data ethics instruction. *Science and Engineering Ethics*, 26(4), 2189-2213.
- ◆ Bezuidenhout, L., & Ratti, E. (2021). What does it mean to embed ethics in data science? An integrative approach based on microethics and virtues. *Ai & Society*, 36(3), 939-953.
- ◆ Calzati, S., & Ploeger, H. (2024). Problem-solving? No, problem-opening! A method to reframe and teach data ethics as a transdisciplinary endeavour. *Big Data & Society*, 11(3), 20539517241270687.
- ◆ Fawns, T. (2022). An entangled pedagogy: Looking beyond the pedagogy–technology dichotomy. *Postdigital Science and Education*, 4(3), 711-728.
- ◆ Knox, J. (2019). What does the ‘postdigital’ mean for education? Three critical perspectives on the digital, with implications for educational research and practice. *Postdigital Science and Education*, 1(2), 357-370.
- ◆ Macgilchrist, F. (2024). Technology, Activism and Living among Planetary Ruins. *Bildung für eine nachhaltige Entwicklung im Umbruch?* <https://library.oapen.org/bitstream/handle/20.500.12657/88626/9783847419068.pdf?sequence=1#page=80>
- ◆ Shi, L. P., & Palenski, T. (2024). From Optimism to Caution: An Analysis of the Seemingly Balanced Policy Discourse on Artificial Intelligence in Education. British Educational Research Association.