fPET 2024 - Book of Abstracts

1) Talks

Agostino Cera	**2A**	DE-METHAPHORIZING TECHNOLOGY: THE QUESTION OF HUMAN-TECHNOLOGY RELATION	My presentation, which deals with the epistemology of both technology and engineering, suggests the necessity of a co-disciplinarity between philosophy of technology and engineering, that is, a form of link whose basic assumption is the acknowledgement that they belong to different epistemic regimes. Specifically, I will focus on a lexical ecology of technology, consisting of a de-metaphorization process. The underlying assumption is that "the (philosophical) question concerning technology" in its current form requires a critical-reflexive work on its language. The traditional lexical trend of technology, the way of telling its extraordinary development, has been that of an anthropologically based metaphorization. Such a choice was justified by the need to make understandable a number of phenomena that would otherwise have remained unintelligible, although their impact on everyone's daily life. The symbol of this traditional trend can be found in Alan Turing's imitation game. Therefore, the classical model of the language and vocabulary of technology, its Turing Stage, was metaphorical or rather mimetic-anthropological. Nowadays, in a Post-Turing Stage, this lexical model is no longer able to perform its orienting function. This crisis depends on a more and more strong approximation of the technical-artificial dimension to the anthropological-natural one, which reaches to the edge of indistinguishability and whose problematic outcomes are self-evident. I refer to the risk that taking these metaphors literally gives birth to a naturalization of technology with inevitable normative fallout. For instance, that artificial intelligence can become the criterion or model to interpret and evaluate human intelligence. Given this premise, the philosophy of technology needs a critical-reflexive work on its lexical dimension – in the form of a de-metaphorization – in order to (re-)emphasize the differences between human and technological features/functions. As a first concrete case of this project, I present that of "relation", namely why human-technology relation cannot be considered a relation. To achieve this goal, I will make use of "the phenomenology of relationality" established by Karl Löwith in his now classic Das Individuum in der Rolle des Mitmenschen (The Individuum in the Role of Fellow Man, 1928). Löwith presents a relational taxonomy which comes from "bond", to "interaction", "connection", "relation" and whose apex is the "responsible relationship (Verhältnis) between an I self and a you self". This paradigm of all forms of relationship can be realized only between "someones", that is between human beings understood as Mitmenschen (fellow men). To make Löwith's argument really functional to my scope, I will supplement it with what I have called the disappointability principle, which is inspired by the "theory of falsification" established by Karl Popper in the Logic of the Scientific Discovery (1934). My point is that to have a genuine Miteinandersein (being-with-one-another), one must be able to be disappointed and disappoint one's co-relatum. As a result, without this requirement a bond cannot be considered an authentic relation/Miteinandersein. Now, insofar as technology is not able, by definition, to disappoint or be disappointed, it is out, by principle, of the relational dimension.
Sarah Junaid	**2A**	Entrepreneurship as an essential vehicle for ethical skills development	Educators in engineering ethics are working hard to instil the necessary knowledge, skills and behaviours to prepare students for the ethical challenges in their profession. Despite these efforts to simulate the workplace context with real-world problems and explore situations students may face at work within the safety of the classroom, developing ethical awareness requires the practical means to build their own experiential knowledge and express these skills and behaviours. Ultimately, consistent ethical behaviours lead to the character traits of an ethically minded engineer. Their ethics education at university, if done well, could be the start of this journey. What has also now become a growing aspect of engineering education over the last decade is addressing the challenges of equality, diversity and inclusion (EDI). EDI as a subset of ethics aims to challenge our assumptions about the social and economic structures in our communities and work collectively towards resolving these. Rather than solving the symptoms of these structures, EDI within engineering looks to address the wider systemic problems. Entrepreneurship is one such way students and young graduates can step into a responsible role that could realistically address these EDI challenges. It has been branded in some university programmes to bring emphasis on the ethical elements of these enterprises such as 'social entrepreneurship' and 'civic entrepreneurship'. Regardless of branding, entrepreneurship is a potentially powerful way to provide the experiential learning needed for these ethical character traits to really develop. This paper explores entrepreneurship in engineering through the lens of ethics education and why we should be doing more to develop ethical awareness through developing entrepreneurship skills.
Yuqi Peng	**2A**	A Rubik's Cube-Inspired Pedagogical Tool for Teaching and Learning Engineering Ethics	To facilitate engineering students' understanding of engineering ethics and support instructors in developing course content, this study introduces an innovative educational tool drawing inspiration from the Rubik's Cube metaphor. The Engineering Ethics Knowledge Rubik's Cube (EEKRC) framework synthesizes six critical aspects of engineering ethics: Ethical Theories lay the philosophical groundwork for ethical analysis, while the Codes of Ethics operationalize these theories into professional guidelines, grounding ethical conduct within engineering. Ethical Issues, encompassing both micro and macro dimensions, facilitate a thorough investigation of ethical dilemmas, whereas Engineering Disciplines address these ethical challenges across specific fields. Stakeholders highlight the impact of engineering decisions on various affected individuals or groups, and the Lifecycle aspect ensures continuous ethical consideration across all project stages. This model, informed by an analysis of engineering ethics textbooks and literature, aims to align with engineering students' learning preferences and organize the broadly referenced yet previously unstructured content cohesively. The Rubik's Cube model is selected for its appeal to engineering students and its capacity to visually and graphically represent complex ethical concepts. This study, focusing on the two-dimensional Rubik's Cube, employs its static layout, indicating that the rearrangement of segments does not affect its classroom application. It outlines the process of constructing the EEKRC, including the motivations for its creation, the selection rationale for the six aspects, and the specific content of each square. It further explores the EEKRC's application in engineering ethics courses, highlighting its utility in aiding instructors with course planning, assisting students in ethical analyses, and serving as an assessment tool for student learning. Additionally, the suitability and potential obstacles of using the Rubik's Cube model to organize content and enhance instructional methods in engineering ethics will be discussed.

Cynthia Pickering, Erik Fisher	**2A**	Cultural Experience as a Source of Ethical Reasoning in STEM Education	A long-standing theme in Engineering Ethics, Philosophy of Technology, and related scholarly discourses, most recently in Responsible Innovation, emphasize the importance of instilling ethical skills and dispositions early on in the training of scientists and engineers. In particular, the literature on Responsible Innovation recognizes that technical practitioners must go beyond standardized ethical codes of conduct and engage in "forward-looking responsibility" (Owen et al. 2012), "Anticipatory Ethical Reasoning" (Conley et al. 2023), and similar types of ethical practice. Anticipatory Ethical Reasoning (AER) occurs when participants apply ethical reasoning that exhibits empathy, fairness, responsibility, character, rights, liberty, personal autonomy, consequences, and/or expectations of legitimate authority to broaden consideration of possible outcomes, and potentially surface unintended consequences. AER is interesting because it goes beyond traditional engineering ethics. This paper reports on an overlooked source of the acquisition of this type of ethical reasoning skill, namely, cultural backgrounds and assets. Specifically, it reports on three Socio-Technical Integration Research (STIR) studies that examined the experience, retention, and advancement of underrepresented minority (URM) students in Science, Technology, Engineering and Mathematics (STEM). Although the focus of these studies was not on ethical reasoning, we nevertheless find that the majority of the 8 participating STEM students engaged in what we call unprompted AER. By unprompted, we mean that, unlike previous forms of cultivating AER in workshop settings, we observed students voluntarily exhibiting AER without having been formally introduced to it. Interestingly, this phenomenon was observed among the 6 underrepresented minority students, whose inherent cultural assets were found to be important aspects of their career and work self-efficacy, but was not observed among the non-URM students. For instance, all three Study 1 participants showed AER about possible outcomes of their own technology decisions on relevant stakeholders in the form of sensitivity, empathy, and fairness. And in Study 3, all three participants showed AER about possible outcomes of their STEM program delivery to relevant stakeholders in the form of sensitivity, empathy, fairness, responsibility, and personal autonomy. By contrast, both Study 2 participants also discussed ethics, but the ethical issues were confined to professional ethics related to the security and privacy of their customers, with little to no instances of AER. These findings potentially indicate that URM students in STEM have inherent cultural assets and alternative knowledges that contribute not only to innovative technology design, development, and implementation, which were the focus of the 3 studies, but also contribute to AER. This presentation will provide qualitative empirical evidence in support of this interpretation. In conclusion, we propose that a key capacity of technical practitioners to engage in Responsible Innovation is to be found in their cultural experiential background as well as in their curricular education. Conley, S. N., Tabas, B., & York, E. (2023). Futures labs: A space for pedagogies of responsible innovation. Journal of Responsible Innovation, 10(1), 2129179. Owen, Richard J., John R. Bessant, and Maggy Heintz, eds. Responsible innovation. Vol. 104. Chichester: Wiley, 2013.
Sarah Carter	**2B**	That's Too Personal: Defining the Limits of Personalization in Mental Health Chatbots	Deploying large language models (LLMs) in healthcare settings raises ethical concerns around autonomy, privacy, and fairness (World Health Organization, 2024). In particular, LLMs utilized for basic talk therapy, often referred to as mental health chatbots, are frequently personalized based on user interactions or other input (Kocaballi et al., 2019). While personalization could improve the patient's experience and quality of care, it could also pose a risk to their autonomy through, for example, the inappropriate use of personalized nudges called "hypernudges" (Yeung, 2017). This presentation aims to define the limits of appropriate, autonomy-preserving personalization in mental health chatbots. To do this, I will use the four-dimensional theory of autonomy (4DT) from Killmister (2017) as a lens to concretely define the myriad of means that autonomy can be promoted and challenged through personalization. 4DT draws on the breadth of autonomy literature to create a comprehensive framework of autonomy, capturing 1.) self-defining (personal identity), 2.) self-realizing (intention and action), 3.) self-unifying (consistency), and 4.) self-constituting (agency) elements. Besides its comprehensiveness, 4DT is especially helpful when defining design parameters for promoting value-centered and autonomous interactions with technology (Carter, 2022). I will firstly consider self-definition by interrogating the effects of chatbot-based personalization on personal identity formation, awareness, and consistency. Next, for self-realization, I will explore personalization-based nudges that encourage certain actions or behavior changes. Similarly, for self-unification, I will analyze the consistency of such actions and changes with the patient's identity and personal values. Lastly, for self-constitution, I will consider how personalization could cause patient overdependence on the system or foster a (renewed) sense of agency. Throughout these dimensions, I will also consider the autonomy ramifications of shifting personal values during a patient's potentially longer-term use of a chatbot system and the challenges introduced by tensions between a patient's personal values. I then will conclude by proposing initial design recommendations based on these insights, including designing personalization based on a patient's personal values to promote reflection and value-consistent actions. Using a 4DT approach, this work contributes added insight into how personalization in chatbots could be used in an autonomy-preserving and appropriate manner to support patient well-being, self-determination, and care. Works Cited Carter, S. E. (2022). A Value-centered Exploration of Data Privacy and Personalized Privacy Assistants. Digital Society, 1(3). https://doi.org/10.1007/s44206-022-00028-w Killmister, S. (2017). Taking the Measure of Autonomy: A Four-Dimensional Theory of Self-Governance (1st ed.). Routledge. https://doi.org/https://doi.org/10.4324/9781315204932 Kocaballi, A. B., Berkovsky, S., Quiroz, J. C., Laranjo, L., Tong, H. L., Rezazadegan, D., Briatore, A., & Coiera, E. (2019). The personalization of conversational agents in health care: Systematic review. In Journal of Medical Internet Research (Vol. 21, Issue 11). JMIR Publications Inc. https://doi.org/10.2196/15360 World Health Organization. (2024). Ethics and governance of artificial intelligence for health: Guidance on large multi-modal models. http://apps.who.int/bookorders. Yeung, K. (2017). 'Hypernudge': Big Data as a mode of regulation by design. Information Communication and Society,20(1), 118–136. https://doi.org/10.1080/1369118X.2016.1186713

Matthias Uhl, Sebastian Krügel	**2B**	Distrust in AI-based decision support systems and ethical implications in medical decision-making	In medical image recognition, AI-based decision support systems can help doctors to quickly identify abnormal structures and patterns. When integrating decision support systems into clinical decision-making processes, however, it is important to consider how human doctors will interact with these systems and how doctors' behavior may be influenced by them. Since empirical research on the behavioral impact of AI-based decision support systems on medical decision-making is currently very limited, we developed an image-based classification task, which can be performed by laypersons, and which entails structural analogies to image-based medical decision-making. Based on this task, we conducted a stream of experiments on human-AI interaction in image-based decision-making. In more than 20 experimental conditions with about 5,000 participants, we investigated the influence of different design factors of human-AI interaction on participants' behavior. We collected about 100,000 image classifications and found a substantial distrust in AI-based decision support systems. This distrust is robust to various information conditions and learning opportunities that we implemented in our experiments. Psychologically, the distrust is rooted in an overestimation of one's own abilities (i.e., overconfidence), an inability of participants to distinguish difficult from simple cases, and an inexcusability in the event of errors by the AI-based decision support system. In this paper, we provide an extensive overview of the experimental results of our studies, put the results into perspective regarding a few studies on AI-based assistance systems with real medical doctors and discuss important ethical implications. One important ethical question relates to the ethical legitimacy of intentionally exploiting the behavioral effects of certain features of the AI-based decision support system that increase doctors' reliance in the system to align their diagnosis more with the AI's recommendations. Put differently, to what extent would it be ethically acceptable to "nudge" doctors to rely more on the AI-based decision support systems by carefully designing the human-AI interaction to achieve this aim? An even more fundamental ethical question raised by our findings may be whether doctors should be taken out of the loop of AI-based diagnostic processes altogether if their involvement merely reduces patients' chances for an accurate diagnosis. How diagnostically dominant does an AI have to be to render our normative insistence on keeping physicians in the loop unethical? And finally, to what extent does a patient have a say when it comes to the design of the interaction between doctors and AI or the respective decision weightings of both agents?
Florian Richter	**2C**	A Philosophical Note on the History of Technology and Progress	Progress is usually intertwined with technology, society, and morality. Technological progress could be seen as a rationalization of means and end connections, i.e., to make certain processes more efficient and effective. Or to provide new options and ends that were not possible before. For example, brewing coffee has become more efficient, and different ways to make coffee have been developed. But drinking coffee was not a possible end for the Ancient Romans. Another technological progress within the field of medicine is the better treatment of toothache. [1, p. 23] However, such technological progress is also embedded in a societal context and connected with health systems in certain regions. Even in richer countries, where many modern methods of medicine are available, people without health insurance suffer because they cannot afford certain treatments. Hence, technological progress is not independent of societal and moral values or economic progress; otherwise. Also, the funding of scientific research and technological development depends on which values are prioritized. Such value decisions also influence the direction of technological progress. It seems that technology is thus something value-neutral because it depends on our interests, needs, and values and how it is used and developed. A gun or even a nuclear missile are just means or tools to protect yourself or to attack someone. [2] However, the big narratives in the philosophy of technology seem to contradict such a position. Some consider technology as something that, in the end, might destroy humanity and culture or as something that might lead to salvation. Of course, the pessimists are, at least in academic debates, more present. For instance, Hans Freyer, Günther Anders, Lewis Mumford, or Jacques Ellul. [3] Furthermore, artificial intelligence leads to fears and worries about the development of some kind of "superintelligence" that might eradicate humanity. [4] On the other hand, transhumanism considers artificial intelligence to be an opportunity to merge with technology [5] or to seek guidance from moral advisors [6]. Technology is thus not value-neutral. Rather, it is either the engine of moral and societal progress or the destroyer of culture and morality. From a postmodern perspective, it is not possible to tell any more such "metanarratives." [7, p. xxiv] But how could we make sense of progress if everything is fragmented into separate narratives or language games? Of course, within certain regions, we might "have less toothache and more mustard, much of it first-rate, from Dijon." Nevertheless, it seems difficult to speak of "Progress" "when seen across two centuries of suffering." However, according to Darnton, "some familiarity with what humanity has suffered in the past may help us appreciate the modest, incremental gains of pleasure over pain or progress with a lowercase p." [1, p. 23] Such a postmodern utilitarianism of looking at progress takes the decrement of suffering and the incremental in pleasure as a standard. Nevertheless, what we consider as suffering and pleasure depends on our framework of values and, thus, also our concept of progress. Should the suffering of animals be avoided or of future generations? Is a society more progressive that does not use animals when testing pharmaceuticals? These are important questions. However, such questions can only be asked and debated in a progressive society. We can set ourselves in relation to values, problematize such issues, and search for legitimizing strategies for positions. [8, p. 141/142] It does not mean to assume a final state of history from where we look back. Such utopian or dystopian visions are usually ideologically imbued with value decisions that are not open for debate. Dogmatic dazzlement or ideological constriction on the societal and moral side prevents the development of autonomous selves that are able to set themselves in relation to values. Technological progress, particularly driven by AI-based and highly automatized systems that are embedded in our life-world, needs to be assessed on such grounds. Is technology enabling us to set ourselves in relation to values, or does it prevent us from developing such competence? The talk will make four contributions. First, the philosophical literature on technological progress is reviewed. Second, the value-neutrality thesis of technology is discussed in relation to progress and the big narratives of the philosophy of technology. Third, Amartya Sen's [9] and Christoph Hubig's [8] conceptual framework for values is outlined in order to find a non-dogmatic and non-ideological foundation for assessing (technological) progress. Particularly, and that is the final aspect of the talk, such a conceptual framework needs to be operationalized within the technological context where artifacts can and may nudge, manipulate, or influence [10] people in their moral judgment. References
Thomas Siller, Gearold Johnson	**2C**	Where does the future end?	The most important concept included in definitions of sustainability is the 'future.' Rarely is a definition of the timeframe of the future provided. It has been suggested that the timeframe extend to billions of years, or the end of the planet's life, others are using 10,000 years as a working timeframe, while much shorter functional timeframes are used for design life of modern technology! Coming to an understanding of the concept of the 'future' is critical to planning for a more sustainable future. A related question is: Can we plan for the future when we get the present so wrong? If we cannot see what is around us, how can we see what is ahead of us? What is needed is a connection between a better understanding of what is, and a philosophical concept of the 'future' that can inform efforts to create a more sustainable society. In this paper, we develope a time-focused ethos to guide us towards a more sustainable future. This ethos should resolve acknowledge the differences between the many timeframes that are associated with technology and future ideas. Looking first at technology we have artifacts that have lifetimes of approximately a year, e.g. cellphones, and others that stretch to many decades, e.g. buildings. A more recently recognized aspect of technological lifetimes is the temporal reach of technology in terms of both material persistence (forever chemicals) and environmental impacts such as climate disruption. These timeframes extend father into the future than the artifacts themselves. Between these two timeframes are the typical human reflections on future generations which often extend only a few past the current generation. We think of children and grandchildren, but father generations are a rarity and therefore less prominent in the human psyche. The ultimate timeframe is the lifetime of the planet, which extends billions of years from the present. The problem with this later concept is that the current trajectory of human behavior makes human extinction it more likely that human life persisting towards the planet's lifetime.

Marc Heimann	**2D**	Bridging Heidegger, Psychoanalysis, and Large Language Models	Abstract: This short talk explores the parallels between the architectural foundations of Large Language Models (LLMs), with a focus on the Transformer model as introduced by Vaswani et al. in 2017, and the rich theoretical underpinnings of hermeneutics and psychoanalysis. It draws on the seminal work of the late Martin Heidegger in hermeneutics and Sigmund Freud and Jacques Lacan in psychoanalysis to build a conceptual bridge between these fields. The Transformer model represents a significant advance in the field of artificial intelligence, particularly in natural language processing (NLP). Its unique architecture, based on the mechanism of self-attention, allows for a more nuanced interpretation of language by evaluating the importance of each word in relation to others within a sentence. This method mirrors Heidegger's hermeneutic circle, a concept that emphasizes the interdependent relationship between the parts and the whole in the process of understanding. Heidegger posited that in order to understand a text (or any object of analysis), one must oscillate between considering its individual components and its whole. This interplay, in the case of primordial words, focuses on the creation of fundamental transcendental horizons through central metaphorical links (Heidegger, 1984, p. 131). Similarly, the Transformer's self-attention mechanism contextualizes data by understanding the meaning of individual words, which can change the entire way the embeddings are used in an interaction, producing effects for which the LLM was not trained, highlighting a parallel process in computational and human understanding. The presentation further explores how this part-and-whole relationship in LLMs echoes the association-based logic found in the psychoanalytic theories of Freud and Lacan. Freud introduced the idea of the unconscious as a repository of latent representational meanings (Freud, 1942, pp. 180-185), structured primarily by linguistic association rather than logical propositions. Lacan extended this view by emphasizing the primacy of language in shaping the unconscious. This perspective resonates with how LLMs, such as those based on the Transformer architecture, use embedding spaces. These are high-dimensional spaces in which words and phrases are represented as vectors, capturing complex relationships and patterns that go beyond surface-level associations. Leveraging these latent embeddings requires prompts that restructure and dynamically link existing patterns into new ones, paralleling Heidegger's primordial words or Lacan's master signifiers. By examining these parallels, the talk seeks to illuminate the interpretive processes that underlie both human cognition and the operation of AI. It argues that the future of AI research could be enriched by incorporating hermeneutic and psychoanalytic theories. For further reading, see Heimann & Hübener 2023. References: Freud, S. (1942). Gesammelte Werke: zweiter und dritter Band. Die Traumdeutung: Über den Traum (A. Freud, Ed.). S. Fischer. Heidegger, M. (1984). Gesamtausgabe: Vol. 45. Grundfragen der Philosophie: Ausgewählte »Probleme« der »Logik« (F.-W. von Hermann, Ed.). Vittorio Klostermann. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention Is All You Need. 31st Conference on Neural Information Processing Systems. Advance online publication. https://doi.org/10.48550/arXiv.1706.03762 Heimann, M., & Hübener, A.-F. (2024). The extimate core of understanding: absolute metaphors, psychosis and large language models. AI & SOCIETY. https://doi.org/10.1007/s00146-024-01971-7
Wolfgang Eppler, Reinhard Heil	**2D**	Towards a Pragmatical Grounding of Large Language Models	Towards a Pragmatical Grounding of Large Language Models Wolfgang Eppler, Reinhard Heil Despite their impressive linguistic performance, LLMs have serious shortcomings such as false information, generalization errors or biases. Linguistic studies help to assess what such models and their further developments are fundamentally capable of. Bender and Koller (2022) have published a noteworthy paper in which they attempt to show that unimodal LLMs neither understand language nor do their linguistic expressions have any meaning. They refer to the symbol grounding problem already described by Harnad (1990), which arises when the elements of a symbol-processing system are exclusively linguistic in nature. If the language elements are not anchored in the world, they have no meaning. Multimodal LLMs have the same problem. Just like texts, images are converted into tokens that have no further reference to the world. Rees (2022) and Gubelmann (2024) raise a justified objection to Bender and Koller. The key terms they use - understanding, meaning and intention - are mental states of actors that have so far been attributed exclusively to humans. There is thus a great danger of denying LLMs mental states simply because they are not human beings. We still think that the grounding problem is valid, but the criticism formulated by Bender and Koller is too strong and therefore easy to attack (Piantadosi and Hill 2022). Our alternative, pragmatist approach, based on speech act theory adapted from Habermas and Brandom's normative pragmatics, does not rely on concepts of mental states, and grounding is seen not as a relationship between linguistic expression and an extra-linguistic event (reference) but as a learning process through practical interaction with things and people. For Habermas (1999), a statement is sufficiently grounded in the objective world if it has proven itself in the lifeworld and can be justified in a discourse with others. Brandom (1994) solved the grounding problem by inference (conclusions of a statement) and anaphora (relations between tokens). Participants in a conversation are giving and asking for reasons and keep score of the arguments communicated. A minimum requirement for that is the memorization of arguments so that they can be referred to, acknowledged or questioned. We show that LLMs lack practice in the world because they are unable to adapt model parameters during dialogs and thus fail to learn from experience. Continuously learning systems that orient themselves in the world could possibly help to compensate for this deficit and improve the trustworthiness of LLMs. References Harnad (1990), The Symbol Grounding Problem, in Physica D 42, 335-346 Bender and Koller (2020), Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data, in Proceedings of the 58th Annual Meeting of the ACL, 5158-5198 Rees (2022), Non-Human Words: On GPT-3 as a Philosophical Laboratory, in Dædalus, the Journal of the American Academy of Arts & Science, 168-182 Gubelmann (2024), Large Language Models, Agency, and Why Speech Acts are Beyond Them (For Now), in Philosophy & Technology, 37 Piantadosi and Hill (2022), Meaning without reference in large language models. arXiv preprint arXiv:2208.02957 Habermas (1999), Wahrheit und Rechtfertigung, 102-137 Brandom (1994), Making it Explicit

Beyza Nur Guler, Qin Zhu	**2D**	Navigating Access: The Impact of AI Recruitment Tools on People with Disabilities in the Workplace	While the minority population is increasing at an unprecedented rate, the disability population remains an untapped source of talent for STEM (science, technology, engineering, and mathematics) education and workforce development. Specifically in STEM, about two-thirds (65%) of workers with at least one disability had less than a bachelor's degree education. Moreover, scientists and engineers with disabilities have a higher unemployment rate than those without disabilities and the overall U.S. unemployment rate in 2019. Overall, 37% of disabled workers maintain their employment for less than 6 months, and 18% keep their jobs for 6-12 months. Despite that an increasingly number of diversity and inclusion initiatives have been created by tech companies, the engineering sector continues to struggle with systemic obstacles that hinder the full participation of disabled individuals. In order to tap into this talent and diversify the engineering workforce, it is especially of importance to understand the barriers encountered in professional settings for engineers with disabilities. This paper addresses the challenges encountered by disabled engineers in the recruitment process, a critical aspect of their professional experience. Throughout this process, discriminatory practices, frequently limit job prospects and hinder career growth for disabled individuals. Additionally, the interview process presents unique challenges, including accessibility barriers, misconceptions regarding capabilities, and biases ingrained in screening algorithms. These challenges are further exacerbated by workplace culture and attitudes towards people with disabilities, resulting in a lack of acceptance, accommodation, and employment opportunities for disabled engineers. This paper explores the intricate challenges disabled engineers encounter in the workplace, particularly within the realm of recruitment practices. Through an extensive literature review, the study aims to address the question: 'What is the impact of AI on recruiting disabled workers?' By synthesizing existing research from databases such as Google Scholar, ERIC, IEEE Xplore, ACM Digital Library, Web of Science, and Scopus, this review seeks to deepen understanding of recruitment practices. The analysis will identify common themes within the literature and discuss their implications, providing valuable insights for researchers, policymakers, and practitioners. Understanding the threats posed by recruitment AI can inform company policies, contribute to the design of fairer AI algorithms, and inform theoretical frameworks and practical strategies for improving recruitment processes.
Martin Stacey	**2E**	On the Cartography of Engineering Knowledge	Although the nature of engineering knowledge has interested philosophers and engineers since the time of Plato and Aristotle, only a few hardy souls have ventured attempts to describe the range of knowledge possessed by engineers. The ones who have are a diverse group who had different purposes and disciplinary perspectives as well as different scopes for the branches of human knowledge they were trying to map: engineering knowledge, design knowledge or technical knowledge. The best known is the aeronautical engineer and historian Walter G. Vincenti, whose 1990 book What Engineers Know and How They Know It sets out to explain to non-engineers what it is that engineers know, in terms that fit how engineers themselves see what they do. Vincenti categorized engineering knowledge by its subject matter rather than its form, into six types: fundamental design concepts, criteria and specifications, theoretical tools, quantitative data, practical considerations, and design instrumentalities. Vincenti has rivals: Wendy Faulkner extended Vincenti's treatment to cover the knowledge used in innovation, coming at the problem from the field of science and technology studies. Faulkner (1994) proposes that two axes cut across the topic area distinctions of technological knowledge. (A) The object of knowledge: the natural world, design practice, experimental R&D, the final product, and knowledge itself. (B) The character of knowledge, along five different subdimensions: Understanding – information – skill, tacit – articulated, complex – simple, local – universal, specific/contingent – general/metalevel. A number of alternative views are rather harder to map to Vincenti, or to each other. The engineer and philosopher Günter Ropohl (1997) proposed a more abstract division into technological laws, functional rules, structural rules, technical know-how and sociotechnical understanding. We agree with Ropohl's contention that sociotechnical understanding is an omission, and an important one, from Vincenti's account. The engineers and design theorists Vladimir Hubka and W. Ernst Eder were concerned with creating a design science, and mapping what design knowledge consists in as a part of that (see Eder, 2008 for a summary of their later views). Besides outlining a variety of areas of the subject matter of design knowledge, they stress a two-dimensional split into descriptive and prescriptive statements, about the products of design and the process of design. Marc de Vries (2003), a philosopher concerned with the nature of technological knowledge, proposed four types of technological knowledge in a more abstract form-of-knowledge oriented typology: One, Functional nature knowledge; two, Physical nature knowledge; three, Means-ends knowledge; and four, Action knowledge. The philosophers Anthonie Meijers and Peter Kroes (2013) were concerned with how engineering knowledge fits a justified-true-belief view of knowledge. They divided it into structural knowledge, functional knowledge, prescriptive knowledge, design knowledge, and know-how; they argue that prescriptive knowledge needs to be accepted rather than believed. In our view none of these typologies are satisfactory. They are too diverse in perspective and purpose for either straightforward combination or opposition. We make some remarks on what a fuller syncretic view might look like, giving due weight to sociotechnical knowledge, knowledge of artefact use, and knowledge of processes including engineering methods. References Eder, W.E. (2008). Theory of technical systems and engineering design science – legacy of Vladimir Hubka. Proceedings of DESIGN 2008. Design Society: Dubrovnik, Croatia. Faulkner, W. (1994). Conceptualising knowledge used in innovation: a second look at the science-technology distinction. Science, Technology, & Human Values, 19(4), 425-458.

Zachary Pirtle	**2E**	What does it mean to do Peer Review at a 'System Level?' Exploring Life Cycle Reviews and the Independent Assessment of Complex Systems Engineering Projects	Systems engineering emerged in the U.S. partially as a response to the increased complexity of engineered systems but also as a way of introducing more control over engineering developments (Johnson 2006). Philosophical and sociological studies of the systems engineering process have explored the role of the systems life cycle as representing good engineering (Akeel and Bell 2013). Life cycle reviews, at NASA and other major systems engineering organizations, often serve as coordination mechanisms to shape work and to develop all relevant portions of a system to a comparable level of maturity by the time of associated life cycle milestones (NASA 1995, 2017). This paper will situate systems engineering LCRs in the context of the history of peer review within science and engineering (Baldwin 2018). The uncertain complexity of engineering projects that are in development, and not yet implemented at scale (and often at great cost), precludes easy peer review of something like 'lab-replicated' results. I will argue that the use of independent assessment teams in LCRs is the closest that systems developers come to true peer review. They allow a multi-disciplinary group of independent reviewers, who generally include engineer 'peers' whom have worked on at least semi-analogous systems, to come in and assess the overall integrated health of an engineering system. Given the holistic and broad complexity of engineering knowledge needed for any given project (Vincenti 1990), LCRs provide a strong snapshot of how all the pieces fit together, which generally cannot be accomplished via other means I will discuss how there are challenges to meaningful peer review at LCRs, which often have independent assessment teams examining in parallel the work of an engineering development team. While there is always hoped to be some generalizable insights to be carried across engineering developments, sometimes the unique context of an engineering effort makes it very difficult to have reviewers from outside the implementing project to be able to fully understand the challenges at hand. Further, the analysis methods used at LCRs, especially those focused on cost and schedule, often invoke a high degree of similarity in method and data input, and can preclude the possibility of a truly independent perspective/cross-check. There can also be pragmatic tradeoffs on the level of depth involved in an LCR, including whether it varies from day-length to multi-week deliberations, which shapes the intensity of discussion and nature of the peer review. I will also discuss, based on public reports from the US Government Accountability Office, how the personality and approach of LCR peer reviewers can strongly shape judgments about a prospective engineering development, and how they need to develop some trust with the team they are assessing in order to establish credibility for their assessments. While there are still cost overruns and enduring critiques of the systems engineering process, the engineering process lives and grows through teams and not through the brilliance of single lone-inventors (Elkins-Tanton 2023). More deeply considering how LCR peer review works in engineering can have an important role in enabling assessment teams to comment on and shape future engineering mega-projects.
Michael Haiden, Florian Richter	**2F**	Autonomous Weapons: Considering the Rights and Interests of Soldiers	Pitt claimed in his seminal paper "Guns Don´t Kill, People Kill" that technology is value-neutral This thesis has been contested by the postphenomenological strand of the philosophy of technology, arguing that tools such as guns embody values (an de Poel & Verbeek 2014, Miller 2021). A layer of complexity is added when autonomous weapons systems come into play. Thus, it is no surprise that scholars increasingly debate the moral consequences of autonomous weapons. Hammond (2015) argues, for example, that accountability of autonomous weapons could become problematic. However, this is not exclusive to autonomous weapons, as states in general are difficult to hold accountable for their conduct in war. Thus, we may be too quick to condemn autonomous weapons. After all, they could limit casualties by replacing soldiers on the battlefield and lower the mental health pressures for drone operatives , which is important since they suffer from post-traumatic stress in roughly equal numbers to airplane pilots (Horowitz, 2016). Furthermore, experts highlight that humans already have a poor record of following the norms of warfare due to emotional reactions in battlefield-situations and that autonomous weapons could be better at following codes of conduct (Amoroso & Tamburrini, 2020). As a result, should we welcome their development if autonomous weapons hardly worsen matters but contain various benefits? This talk aims to examine this question. Our focus will be on the rights and interests of soldiers. Given that those employed in the military suffer from physical and mental pressures, autonomous weapons would be a positive development if they prevent such problems and do not carry counterweighing drawbacks. We argue that the benefits for soldiers are easy to discern – especially in countries where they are involuntarily drafted – while the negative sides of autonomous weapons are more speculative and may not create new moral issues in warfare. From this perspective, there is a case in favor of autonomous weapons. References Amoroso, D., & Tamburrini, G. (2020). Autonomous Weapons Systems and Meaningful Human Control: Ethical and Legal Issues. Current Robotics Reports, 1(4), 187–194. https://doi.org/10.1007/s43154-020-00024-3 Hammond, D. (2015). Autonomous Weapons and the Problem of State Accountability. Chicago Journal of International Law, 15(2), pp. 652–687. Horowitz, M. C. (2016). The Ethics & Morality of Robotic Warfare: Assessing the Debate over Autonomous Weapons. Daedalus, 145(4), 25–36. Miller, B. (2021). Is technology value-neutral? Science, Technology, & Human Values, 46(1), pp. 53-80. Pitt, J. C. (2014). "Guns Don't Kill, People Kill"; Values in and/or Around Technologies. In Kroes, Peter, & P.-P. Verbeek (Eds.), The Moral Status of Technical Artefacts. Philosophy of Engineering and Technology (Vol. 17, pp. 89–101). Dordrecht. van de Poel, I., & Verbeek, P.-P. (2014). Can technology embody values? In P. Kroes, & P.-P. Verbeek (Eds.), The Moral Status of Technical Artefacts. Philosophy of Engineering and Technology (Vol. 17, pp. 103-124). Heidelberg/New York/London.
Nathan Wood	**2F**	Explainable AI in the Military Domain	Artificial intelligence (AI) has become nearly ubiquitous in modern society, from components of mobile applications to medical support systems, and everything in between. In societally impactful systems imbued with AI, there has been increasing concern related to opaque AI, that is, artificial intelligence where it is unclear how or why certain decisions are reached. This has led to a recent boom in research on "explainable AI" (XAI), or approaches to making AI more explainable and understandable to human users. In the military domain, numerous bodies have argued that autonomous and AI-enabled weapon systems ought not incorporate unexplainable AI, with the International Committee of the Red Cross and the United States Department of Defense both explicitly including explainability as a relevant factor in the development and use of such systems. In this article, I present a cautiously critical assessment of this view, arguing that explainability will be irrelevant for many current and near-future autonomous systems in the military (which do not incorporate any AI), that it will be trivially incorporated into most military systems which do possess AI (as these generally possess simpler AI systems), and that for those systems with genuinely opaque AI, explainability will prove to be a more limited value than one might imagine. In particular, I argue that explainability, while indeed being a virtue in design, is a virtue aimed primarily at designers and troubleshooters of AI-enabled systems, but is far less relevant for users and handlers actually deploying these systems. I further argue that human-machine teaming is a far more important element of responsibly using AI for military purposes, adding that explainability may undermine efforts to improve human-machine teamings by creating a prima facie sense that the AI, due to its explainability, may be utilized with little (or less) potential for mistakes. I conclude by clarifying that the arguments are not against XAI in the military, but are instead intended as a caution against over-inflating the value of XAI in this domain, or ignoring the limitations and potential pitfalls of this approach.

Martin Stacey, Claudia Eckert	**3A**	Engineering Knowledge and Soft Skills	The question of what is engineering knowledge is of great interest to philosophers of technology as well as to engineering academics, who have to teach and assess engineering knowledge. Both groups look at this largely from the perspective of individual engineers and what they know or need to know. However, engineering is also a sociotechnical activity: it is almost always collaborative, involving social processes at different scales. It is carried out by teams in organisations and across complex supply chains. Engineers often need to interact with colleagues with different disciplinary backgrounds, and consider the perspectives of users in different contexts. The life cycles of most products are deeply influenced by their use contexts, which are often highly varied, as different users deploy products in intended and unintended ways. Moreover, different stakeholders in the use of technical systems often have different perspectives and timescales. Designers and other technical workers need to understand how different users use or misuse a system, or how colleagues from a different field conceptualise a problem. To work successfully, engineers don't just need technical knowledge and knowledge of how to design. They also need what are termed 'soft skills', both for problem solving and for communicating ideas and information and handling the personal interactions involved in technical work. This requires general tacit and explicit knowledge about human interactions, but also an understanding of the engineering knowledge that other parties bring to collaborative engineering activities. The need for 'soft skills' in engineering teams and other technical professionals raises a question for educators and managers: How can valuable abilities be taught or trained, besides giving people opportunities to practise them? This in turn raises the question of what we want to teach or develop. 'Soft skills' is a loose term, but one that covers a number of identifiable clusters of abilities. We consider some of these, to ask whether they contain aspects of what can usefully be considered knowledge, and if so, how far the knowledge is domain-specific, and if so, how far engineering or technical knowledge is distinctive. We consider how the knowledge involved in key soft skills fits into the broader spectrum of engineering knowledge, by relating it to existing classifications of engineering knowledge produced by different authors for different purposes.
Yafeng Wang	**3A**	Criteria of Success for Engineering Accident Investigations: A Question-Centered Account	Engineering accident investigations are systematic inquiries into the facts and causes of engineering accidents. These investigations serve many important aims, including identifying what happened during the accidents, learning effective lessons to improve safety, helping those who were adversely affected attain closure, and having a calming effect on the public. Consequently, many individuals and groups need to evaluate how successful an engineering accident investigation is in achieving these aims. In this paper, I propose normative criteria for evaluating the degree of success of an engineering accident investigation using conceptual resources drawn from epistemology and philosophy of science. The basic idea of my proposal is that an engineering accident investigation is successful to the extent that (1) questions that should arise in the investigation do arise, and (2) questions that arise—especially the more significant ones—are resolved satisfactorily by the end of the investigation. My paper unpacks this proposal by analyzing the following three concepts and illustrating them using examples from the TWA Flight 800 accident investigation: The (satisfactory) resolution of questions, the significance of questions, and the arising of questions. First, I analyze question resolution in terms of a cognitive attitude called "full and official acceptance". Second, I distinguish four types of question significance in engineering accident investigations: Intellectual, practical, instrumental, and critical significance. Third, I conceptualize the arising of a question as a change in the investigators' cognitive attitude from not deeming the question worthy of pursuit to deeming it worthy of pursuit.
Michael Poznic, Vivek Kant	**3A**	Structuring of information for Human Machine Interaction: Engineering understanding, complex sociotechnical systems, and interface design	While digitalization in the technological sector grows in various industries throughout the world, there is a need to comprehend how humans interact with this digital technology. Specifically, an area of engineering known as Human Machine Interaction endeavors to design interfaces in complex sociotechnical systems. The aim of this paper is to analyze what is required for interface design in the context of such complex systems from a philosophical point of view. The main perspective is an epistemological one and it focuses on engineering understanding as the epistemic achievement interface designers are striving for. So, the topic of this paper is understanding interface design. We will discuss a particular example of an interface within an energy infrastructure. The interface we're dealing with is actually an ensemble of interfaces that is built into the control room of a power plant. At the heart of the topic is the interface as a relation between the human and a machine. The dynamics of the machine is different from the dynamics of the human user, yet, an interface links these two disparate entities. A specific manner in which the interface connects these entities is through the structuring of information. This structuring is organized in two ways, mainly: spatially and temporally. This structuring of information enables the human user to comprehend the information and to act on it. However, the structuring of information is itself not enough as it depends on a number of other factors that are beyond those of immediate operations. These include constraints posed by the activities of other team members, organizational policies, as well as issues of standards and regulations. These additional factors are extra-individual factors that the user does not necessarily has to be fully aware of. However, some of them are needed for the user's handling of the machine. The designer has to take all of these extra-individual factors into account, though. When the interface is designed, it ultimately depends on these individual and extra-individual factors. The individual factors are important for the user in order to deal with specific problems of managing the functioning of the machine. The challenge for the designer is to gain an understanding of how these different aspects of the problems come together as reflected in the individual and extra-individual factors. There are various forms of knowledge the engineering designer has to compile. Yet, how they need to be collated and comprehended together to yield the understanding required to design an interface for a complex sociotechnical system is an open question. Some of the individual factors are not needed for the designer to understand the workings of the interface sufficiently. We will discuss a concrete example of an interface design to spell out the different roles of the factors as regards the users and designers of such interfaces.
Zhang Tongkuo	**3B**	Reflection on the Ethical Issues of Brain Computer Interface from the Perspective of Responsibility Ethics	Abstract: Unlike traditional technological practices, Brain-Computer Interface (BCI) technology has transcended the old metaphysical paradigm of the "man versus technology" dichotomy, achieving a deep integration between humans and technology. From the perspective of responsibility ethics, as BCI technology becomes widely applied in fields such as rehabilitation and entertainment, ethical risks such as privacy paradoxes, "technological cocoons," identity dilemmas, and responsibility disputes are gradually emerging. These potential ethical risks require analysis beyond the limitations of traditional ethics, which often focus on the immediacy of responsibility, and instead call for a broader view of responsibility ethics. According to the principles of responsibility ethics, there is a need to adopt a forward-looking and preventive perspective to examine the motivations of the principal entities, objects, and stakeholders involved in BCI technology. Based on this understanding, it is proposed to incorporate principles of autonomy, goodness, trust, and fairness into forward-looking responsibility ethics research. By adhering to ethical principles such as responsibility attribution, technological transparency, technological trust, and fairness, it is possible to foster a deeper respect for technology and regulate the behavior of those involved. Furthermore, it is essential to strengthen dialogue between researchers and the public, tailored to the needs of different stakeholders, in order to overcome the challenges of traditional responsibility ethics and prevent potential harm to humans, future generations, society, and future societies.

Adam Briggle	**3B**	Evaluating the Ethics and Politics of Gender Affirming Care for Trans Youth	Pediatric medicine is an important site for the analysis of the ethical and political dimensions of technology. Every day, millions of families and health care practitioners wrestle with difficult decisions about how to care for children, which often necessitates weighing the risks and benefits of various medicines and techniques. Over the past five years, gender affirming care for transgender youth has been politicized and has become more controversial than perhaps any other field in pediatric medicine. This presentation utilizes tools from bioethics and STS to assess the controversy. It argues that, contrary to increasing political efforts to restrict gender affirming care, the ethical path forward is greater access to individualized, holistic care for gender non-conforming youth. The first part of the presentation will provide social, political, and historical context. It will draw from Julian Gill-Peterson's Histories of the Transgender Child (2018), Susan Stryker's Transgender History (2017), Jack Turban's Free to Be (2024), and other sources to define key terms and situate the analysis. It will then survey the politicization of gender affirming care with an emphasis on the U.S. context, but with reference to European nations and the U.K. (e.g., the recently released Cass report). The second part of the presentation will ask: is gender affirming care for transgender youth ethical? It will argue in the affirmative through an analysis rooted in the standard bioethics framework of beneficence, non-maleficence, autonomy, and justice. Issues to be explored include the appropriate standards of evidence (e.g., the principle of equipoise in deciding when randomized control trials are ethical), ethical and epistemic risk, regret rates, the absence of any neutral choice, and the notion of an open future. This analysis culminates in the conclusion that justice calls for greater, not less, access to care. The third part of the presentation will attempt to diagnose the root causes of the politicization of gender affirming care for trans youth. In other words, given that there isn't anything different about the ethics of gender affirming care vs other forms of pediatric medicine, why has it been subject to so much scrutiny, vitriol, and criminalization? An answer will be offered drawing from Julia Serano's notion of "gender entitlement" to highlight the double standard applied to transgender vs cisgender people. Indeed, the very same medicines and techniques targeted in the case of trans youth are readily accepted in the case of cisgender youth. In this way, the presentation will argue that transphobia, rather than any legitimate ethical concern, underwrites the political attacks on gender affirming care for trans youth.
Ozan Gurcan	**3B**	Equality and Reprogenetic Autonomy in the Genomic Era	In this paper I question the argument from human dignity in the Universal Declaration on the Human Genome and Human Rights (UDHGHR) and in the recent views of the International Bioethics Committee (IBC). I focus on what this argument says about the permissibility of reprogenetic choices that may be available to prospective parents in the genomic era. The argument from human dignity holds that non-medical genetic choices and enhancements ought to be prohibited because they violate the principle human dignity. I argue that the principle of human dignity/equality need not be violated by the enterprise of human genetic selection/enhancement if reasonable social safeguards are established. In particular, I argue that respecting the reprogenetic autonomy of the decision-maker is paramount within the boundaries of (i) prohibiting the infliction of a shortened lifespan or pain upon the child; (ii) prohibiting the actualization of demeaning beliefs or intentions such as viewing certain groups as inferior; and (iii) prohibiting the choice resulting from an expression of unwillingness to love and care for the child. With these limits, choices cannot be said to undermine equality or dignity by creating unjustified harms or expressing demeaning ideas.
Gearold Johnson, Thomas Siller	**3C**	Utopian Visions: A Critique of Mega City 2070	Utopias as communities or societies (imaginary or in some cases, actual) offering highly desirable or perfect qualities have existed for over two thousand years. They have been envisioned for addressing political, economic and religious issues. In almost all cases, any attempt to create a utopian society has resulted in the opposite result. Engineers and technologists are not immune from their own utopian visions as evidenced from exhibits of tomorrow's products at world's fairs, futures-based theme parks and the just recently announced group of tech billionaires trying to create a utopian city east of San Francisco, U.S.A. Now the American Society of Civil Engineers (ASCE) has developed a large augmented/virtual reality (A/VR) simulation of a Mega City in the year 2070 with a population of 50 million (www.futureworldvision.org). An IMAX film version of this model, "Cities of the Future" was released in science museums around the world with much fanfare on February 16th of this year. The primary reason given for this project is to show how exciting the future will be for young people in pursuing civil engineering studies. Such a goal is noble, but serious questions about the model remain. First, why is the future defined as 2070 when the United Nations predicts that global population will peak at 10.5 million? Many groups also predict that near that date the Earth's population will begin its precipitous decline undoing the exponential growth of the previous two and one-half centuries. In this scenario, mega cities might face far more significant degrowth issues than presented by ASCE. Another issue is that most mega cities, if they develop, will be located on coasts with no discussion of climate change disruptions like sea-level rise. The paper will present a critique of Mega City 2070 including serious demographic deficiencies, a complete lack of human social services and support, politico-economic issues, and omission of key environmental and even technological concerns. The bottom line is that Mega City 2070 consists only of systematic technological innovations and interventions as a logical extension of the past applied city-wide, simultaneously, in stark contrast to how "real" cities evolve. It also ignores the 'end of stationarity' which results in the future based on possible discontinuities from the past. And finally, a discussion on how to incorporate Mega City 2070 into current civil engineering education will be presented.

Katherine Goodman	**3C**	The good of engineering: from a current state to a preferred one	If the ideal good from the practice of law is justice and the ideal good from the practice of medicine is health (Mitcham), then one might argue that the ideal good from the practice of engineering is progress. While progress is a word with many, varied connotations, both positive and negative, the word is chosen deliberately, to draw on both those contradictory connotations and the notion of 'change deliberately shaped for human preferences.' The world is constantly changing, by deliberate human action, through unintended consequences of human action, and from non-human natural forces. The aim of engineering, then, is to shape change with specific intent. While this may seem overly broad, we understand engineering to be a human activity that predates the coining of the word engineering (Koen) and encompasses both detailed planned work and acts of bricolage or tinkering to accomplish its aims (Johnson). Engineering design allows humans do things with higher performance or enhanced capabilities of some kind. In fact, one humble definition of design is the work to move from a "current state to a preferred state" (Zimmerman et al.). This raises the question: Preferred by whom? Which humans benefit from these changes? In ideal circumstances, the general populace would be the answer, with careful planning to avoid the most harmful side effects. Yet most engineering projects are directed by economic interests, often with an aim for profit. Individual engineers typically do not choose the nature of the projects they work on or the requirements shaping that work. If institutions select the projects, and therefore the "preferred state," this implies that those institutions are the moral agents, the entities that determine what good might be attained through their efforts. This is an ethically complex situation for individual engineers to navigate (Goldberg), but it does not remove their ethical obligations. This paper proposes to explore the notion of engineering's good as progress, outlining the insights it may provide and the limitations it imposes. Goldberg, David E. "Is Engineering Philosophically Weak?" Philosophy and Engineering: Reflections on Practice, Principles and Process, edited by Diane P Michelfelder et al., Springer Netherlands, 2013, pp. 391–405. Springer Link, https://doi.org/10.1007/978-94-007-7762-0_30. Johnson, Christopher. "Bricoleur and Bricolage: From Metaphor to Universal Concept." Paragraph, vol. 35, no. 3, Nov. 2012, pp. 355–72. Edinburgh University Press Journals, https://doi.org/10.3366/para.2012.0064. Koen, Billy Vaughn. "Debunking Contemporary Myths Concerning Engineering." Philosophy and Engineering: Reflections on Practice, Principles and Process, edited by Diane P Michelfelder et al., Springer Netherlands, 2013, pp. 115–37. Springer Link, https://doi.org/10.1007/978-94-007-7762-0_10. Mitcham, Carl. "A Philosophical Inadequacy of Engineering:" Monist, edited by Sherwood J. B. Sugden, vol. 92, no. 3, 2009, pp. 339–56. DOI.org (Crossref), https://doi.org/10.5840/monist200992320. Zimmerman, John, et al. "Research through Design as a Method for Interaction Design Research in HCI." Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, 2007, pp. 493–502. DOI.org (Crossref), https://doi.org/10.1145/1240624.1240704.
Xue Yu	**3C**	The Emerging Forms of Human-Machine Relation and Its Philosophical Reconstruction in the Era of Deep Intelligence	The human-machine relationship has always been a core issue of concern in the philosophy of technology. The era of deep intelligence has brought about new changes in human-machine relationships, triggering various emerging forms of human-machine relationships. Based on current technological progress, the emerging forms of human-machine relationship mainly manifest in four aspects: interface relationships based on human-machine interaction, connection relationships based on human-machine networks, action relationships based on human-machine cooperation, and fusion relationships based on human-machine systems. In the comprehensive integration process of these four emerging forms of human-machine, some new features have emerged, such as interactivity, heterogeneity, collaboration, and fusion. Tracing back to the origin, the emerging forms and emerging characteristics between humans and machines suggest isomorphic relationships between them. Generally speaking, humans and machines have the same or similar system structures, which are manifested in similar internal relationships such as action logic, cognitive patterns, and social functions. On the one hand, this is because humans project their own functions and structures onto machines, and on the other hand, it is also the reverse shaping of humans by machines. This inherent relationship of mutual constitutive relationships is also a deeper traceability of human-machine isomorphism. In the era of deep intelligence, the isomorphic relationship between humans and machines is directly manifested as a two-way rush between the humanization of machines and the mechanization of humans. As machines become more and more like humans, humans also exhibit some machine features, such as fragmentation and automation. This article attempts to propose an explanatory path and future vision for human-machine symbiosis in the face of the emerging forms and practical problems of human-machine in the era of deep intelligence. The term "symbiosis" first appeared in the field of biology, emphasizing the close connection between different organisms. Inspired by the typological metaphor of biological symbiosis theory, human-machine symbiosis can be divided into three modes: biased symbiosis, biased symbiosis, and mutually beneficial symbiosis. These three modes respectively demonstrate the cooperation and competition relationship between humans and machines in different interaction scenarios. Examining the emerging forms and characteristics of human-machine symbiosis within the framework, and analyzing the coexistence, dependence, and transformation of benefits and harms, may provide a new explanatory perspective for the current evolution of human-machine relationships.

Johannes Brinz	**3D**	Neurormophic AI: From Simulating Towards Replicating the Brain	With AI systems increasingly penetrating our every day life world, one of the main objectives of politics and society is to evaluate its potential risks and benefits. One aspect that starts to play a role is that of artificial consciousness. While there are cases in which it might be particularly useful to have sentient artificial agents, think of nursing robots in health and elderly care for example, there are other areas in which we want to use AI systems as tools that we can dispose of at our will. In law and ethics there seems to be a widespread consensus that artificial agents should be granted some kind of moral or legal status as soon as they develop (valenced) conscious experience, i.e. the capability to feel pleasure and pain. Some even think that the development of conscious AI systems might lead to a "suffering explosion" and that politics should ban research that risks the emergence of artificial consciousness until we have a better understanding of the matter. While digital computers have been the state of the art in artificial intelligence for at least the last seven decades, a paradigm shift seems to be imminent. With Moores law coming to an end and due to limitations regarding speed and energy efficiency, computer engineering has turned its attention towards computing systems beyond the von Neumann architecture. Of particular interest are chips that work with mechanisms analogous to those in the brain. So-called neuromorphic hardware consists of interconnected physical "neurons" and "synapses", and promises to reduce the power consumption of large AI models by two orders of magnitude, thus potentially replacing digital AI implementations within the next few decades. Some neuromorphic chips are developed by profit-oriented technology corporations such as the Intel or IBM. The question of whether neuromorphic features bias advanced AI towards consciousness remains an open question. The present paper intents to make up for this shortcoming. I argue that implementations of AI algorithms on brain-like computers are more likely to generate artificial consciousness than simulations on standard digital hardware. First, I give a working definition of what a model is. Based on that definition, I distinguish between simulations and replications. A simulation only shares a common mathematical structure with the model it simulates, whereas a replication additionally must be the sort of entity the model is about (assignment) and work according to the causal laws described by the model (explanation). Then I argue that a property is organisationally invariant if and only if it is due to the abstract causal structure. In contrast to that, other non invariant properties depend on the replicating causal structure that is shared by all replications of the same model but no by (all) its simulations. Consciousness is taken to be potentially not organizational invariant. The last section then explores neuromorphic hardware. I argue that neuromorphic chips share parts of the replicating causal structure of biological brains, and are, thus, more likely to instantiate the mechanisms underlying consciousness.
Beatrice Bonami	**3D**	Social Organic Authenticity and non-Western Southern Epistemology as a Core Pathway for Future Artificial Intelligence	In Science Fiction's history, narratives have depicted intelligent systems as a subservient class capable of gaining power over humans, often positioning machines as entities distinct from organic life, suggesting their inability to achieve authenticity or consciousness. Underlying these narratives is an apprehension that machines may surpass humans in intelligence or seek retribution for historical subjugation (Foster, Burkett, 2000) reflecting concerns about the potential autonomy and demands for justice from inorganic forms of intelligence. The boundaries between natural and crafted beings are blurred and anticipate a future where humans merge with technology, challenging traditional notions of kinship (Haraway, 1988, 1991) and highlighting the emergence of new social structures in a bimodal society. The question of what defines humanity against machines becomes increasingly complex, particularly concerning authenticity (Bourdieu, 1977, Corcoran, 2009). While machines may emulate aspects of efficiency, they lack authentic human organicity (Mancuso, 2020; Doherty, 2012), although recent advancements in AI pose new challenges to traditional understandings of human uniqueness. Addressing these tensions requires a nuanced approach considering diverse perspectives, particularly from non-Western cultures. By examining African and South American literature (Mignolo, 2007; Escobar, 2017; Viveiros de Castro, 2015; Cordeiro, 2023; Mbigi, 1997; Mugumbate, Nyanguru, 2013; Ngwa, N. et al. 2023; Haenen, 2013), we can challenge Eurocentric imaginaries of AI and explore alternative frameworks for its integration into future society. By incorporating indigenous and post-colonial knowledge, this paper aims to contribute to this discourse by supplementing the global dialogue on AI ethics and governance. Bourdieu, P. (1977) Outline of a theory of practice (R. Nice, Trans.). Cambridge: Cambridge University Press. Corcoran, T. (2009) Second nature. British Journal of Social Psychology. Cordeiro, A. (2023) Pajé No Se Transforma, Se Ingera: La Relación Entre Bicho y Gente. Punto Rojo Editorial: Madrid, Spain. Doherty, H. (2012) Organic Philosophy (Volume 1); Or, Man's True Place in Nature Epicosmology. General Books. Escobar, A. (2017). Designs for the Pluriverse: Radical Interdependence, Autonomy, and the Making of Worlds. Duke University Press. Foster, B., Burkett, P. (2000). The dialectic of organic/inorganic relations. Organization & Environment. Haenen, H. (2013) One person cannot embrace a Baobab. Leiden, The Netherlands: Brill. Haraway, D. (1988) Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. Feminist Studies. Haraway, D. (1991) Cyborg Manifesto: Science, Technology, and Socialist-Feminism. New York: Routledge. Mancuso, S. (2020) The Incredible Journey of Plants. Other Press. Mbigi, L. (1997) Ubuntu: the African dream in management. Randburg: Knowledge Resources. Mignolo, W. (2007) El pensamiento decolonial: desprendimiento y apertura: Un manifiesto. In S. CastroGómez and R. Grosfoguel (Eds.) Bogotá: Siglo del Hombre Editores. Mugumbate, J. Nyanguru, A. (2013) Exploring African philosophy. African Journal of Social Work. Ngwa, N. et al. (2023) Life under the Baobab Tree. Transdisciplinary Theological Colloquia. Viveiros de Castro, E. (2015) Metafísicas canibais: elementos para uma antropologia estrutural. São Paulo: Cosac Naify.

Arzu Formanek, Robert Miehe, Klaus Erlach, Yannick Baumgarten	**3D**	Are Biointelligent Systems Intelligent? The Onion Model for Biointelligence	Emerging Biointelligent Technologies promise a path to achieve sustainability goals regarding our technological coexistence with the natural environment. There are attempts to define the specifics of Biointelligent Manufacturing (BIM)[1] in recent literature. BIM systems are based on systemic complex interactions of technical, biological and information technological components (i.e. bio-, hard- and software, especially AI systems). They are designed to enable a highly localized on-demand production of personalized goods via stand-alone non-expert systems. The normative goal is to aim at a sustainable balance with the environment by achieving flexibility, efficiency and autonomy (following nature's example). All these and many similar characteristics are what underly the reference to 'intelligence' in the technical notion of 'biointelligence'. The question however remains: in what sense, if any, are BIM systems intelligent? As there is no conceptually mature and shared concept of 'intelligence'[2], addressing such a question proves to be a challenging task. Thus, to answer this question we develop a structured and layered model of how we extend intelligence to biointelligence: the onion model. This model has three main layers which go from skin to germ. The skin contains the most mature and readily agreed upon extension of intelligence: the human technoscientific endeavor. This is in line with the general idea[3] that humans technologically shape their environment to increase their potentials of intelligence. But, since this could apply to many technologies that can share the sustainability goal in BIM, this layer is peeled away easily. It reveals the middle layer which contains intelligence as deployed by biointelligent manufacturing ecosystems. The middle layer in turn gives way to the germ, containing intelligence as deployed in individual biointelligent technologies, i.e. products. Referring to specific characteristics of BIM systems within these layers, we compare and discuss several positions on intelligence from relevant fields of philosophy[4]. We show that these positions do not capture biointelligence, as they fail to account for novel systems like BIM that systematically combine biological and information technological components with a normative goal of achieving environmental balance. Structurally unfolding all these characteristics, the onion model allows us to show in which practically, conceptually, and techno-socially informative ways BIM systems are intelligent. [1] Miehe, R., Buckreus, L., Kiemel, S., Sauer, A., & Bauernhansl, T. (2021). A Conceptual Framework for Biointelligent Production—Calling for Systemic Life Cycle Thinking in Cellular Units. Clean Technologies, 3(4), Article 4. [2] Legg, S., & Hutter, M. (2007). A Collection of Definitions of Intelligence. https://doi.org/10.48550/arXiv.0706.3639 [3] see also: Müller, V. C. (2024). Philosophy of Ai: A Structured Overview. In N. A. Smuha (Ed.), Cambridge handbook on the law, ethics and policy of Artificial Intelligence (pp. 1–25). Cambridge University Press. [4] For example: Coelho Mollo, D. (2022). Intelligent Behaviour. Erkenntnis. Cruse, H., Dean, J., & Ritter, H. (Eds.). (2000). Prerational Intelligence: Adaptive Behavior and Intelligent Systems Without Symbols and Logic, Volume 1, Volume 2 Prerational Intelligence: Interdisciplinary Perspectives on the Behavior of Natural and Artificial Systems, Volume 3 (Vol. 26). Springer Netherlands.
Klaus Erlach, Thomas Bauernhansl	**3E**	How to Formulate a Research Question in Applied Engineering Science: A Systematic Approach	Getting started with a research question in applied engineering science is a challenge in a twofold manner. On the one hand, research findings should be applicable in technological practice, rather than just presenting a theoretical outcome, which may not immediately contribute to practical technological usage. In short: practical validation of findings is a prerequisite for their relevance. On the other hand, research results should be generalizable to more than only a single application, and should not result in just a successful development of one fancy gadget. In short: formal verification of results is a prerequisite for their correctness. Especially in setting up a PhD-thesis it is important to formulate the research question adequately in this respect. Methodically improving this process significantly empowers technological education. For that reason, we have developed a systematic approach for how to formulate efficient research questions in applied engineering science. Our approach consists of dividing the formulation of research question into two steps: a guiding question and up to four types of sub-questions, which correspond to different categories of scientific goals. These questions are formulated using the question word "how"- like any technological question (cf. Blumenberg). The problems are the starting point, which are targeted to be solved by application of the intended research results. Therefore, the guiding question should address the specific features of the to-be-designed new technical or methodical system–in the following called "technosystem". The guiding question asks for the general system design, thus its character is to be the main design question. Here the technosystem must have a dedicated purpose and is to be developed for selected adopter in a delimited context. Thus, our formula for the design question is as follows: "How can a <target-oriented by features F1, F2 …> technical respectively methodical system for fulfilment of <purpose P> for <selected adopters A> in <delimited context C> be designed?" Role of the sub-questions is to clarify which specific research work has to be done to reach the objective of guiding question. For that, we apply the differentiation of three generic categories of scientific goals: the descriptive goal, the theoretical goal, and the pragmatic goal of applied science (cf. Töpfer, Chmielewicz). Each of them must be addressed by one sub-question at least – otherwise one central property of applied engineering science would be left out. First, the descriptive goal of applied science is to find the answer to the modelling question. This includes several issues from concept formation, justified selection of modelling language, modelling of the research scope with its pragmatic orientation (cf. Stachowiak) or building up a reference model. The question here to be answered is: "How to model the scope of research?" Second, with respect to the theoretical goal of applied science, two fundamental research approaches have to be distinguished with respect to their specific ways of questioning. First of these approaches is to find the answer to the explanation question, which means to derive nomological relations between causes and effects, typically explored by experiments. In this case, the question to be answered is: "How can technical effects be explained by regularities likewise causality?" The second approach is to find the answer to the configuration question, which means to derive design principles and design guidelines as teleological linkage for configuration of the target state of a technosystem (cf. Erlach). The so-called Design Science Research presents some useful procedures (cf. Peffers, Holmström). However, there isn't a mature methodological theory for this kind of approach yet. In this case, the question to be answered is: "How can technosystems be configured by teleological guidelines of design?" Third, the pragmatic goal of applied science is to find an answer to the implementation question, which prescribes the procedure to apply the research findings in industrial or technical practise. Here we find the steps of well-known as well as new developed rule-based activities, which lead to the intended result. This last question answers the relevance of the research results. The question to be answered is: "How to proceed in designing the new or optimized technosystem?"

Daniel Marom	**3E**	To educate is not to engineer - why it is important to make the distinction in the education of engineers	Paper Proposal: To educate is not to engineer – why it is important to make the philosophical distinction in the education of engineers In 1985, the renowned philosopher of science and of education, Israel Scheffler, addressed the question "Computers at School?" in a distinguished lecture series at the Interactive Technology in Education Program at the Harvard Graduate School of Education (Scheffler, 1986). In his lecture, Scheffler warned against the narrowing effect the hypnotic fascination with computers could have in the determination of the aims and scope of education. "[A]s the computer's presence grows, the whole array of our educational ends tends to shrink to what is supposedly achievable by a computer." Scheffler also warned about the corrosive effects metaphoric language in digital culture could have on teaching and learning by conceiving of them exclusively in terms of the information processing with the learner's brain becoming a kind of computer. Almost four decades later, it is hard not to relate to Scheffler's talk as prophetic. Alongside the benefits of computer technology and digital culture in education, the educational, psychological, and social damage has been significant (Turkle, 2015, Haidt, 2023). The damage has not only been in sidelining educational aims such as the pursuit of aims such as the development of character, cultural induction, and social and civic engagement. It has also been in the domain of cognitive development. Attention span on any one topic has dwindled to 49-seconds, limiting reading and writing capacities of learners and hindering the development of the brain's executive function (Mark, 2023) The proposed paper will argue this damage is paradoxically most consequential in its impact on the field that makes computer technology and digital culture possible – the education of engineers. Without overcoming the corrosive effects engineering has had on education in general, engineering education will be unable to equip the next generations of engineers with the dispositions, modes of thinking and competencies that are necessary for them to be equal to the daunting and complex tasks that technological innovation itself has generated. This argument is demonstrated through philosophical analysis of the curricular foundations and pedagogical modes persistently employed in mainstream training of engineers. It is argued that this kind of educating proactively contributes to what the philosopher of science Stephen Goldman calls "the social captivity of engineering" by reducing its ends to service provision rather than expanding them to empowering the identity and agency of the engineering profession. Even current change efforts in engineering education such as the alignment of teaching with evidence-based research from the learning sciences, the integration of ethics, teaching systems thinking, and cultivating engineering leadership are overly guided by an engineering mindset to education rather than one that is educational in its essence. It will be argued that to succeed in educating engineers, systematic efforts must instead be made in 1) transforming engineering education through the articulation of visions of engineering education; 2) creating transitional learning processes and rites of passage at the level of induction, graduation, lifelong learning; 3) training leaders, faculty instructors and in-service learning facilitators (in distinction from "mentors") to be engineering educators – a role that is different than that of being an engineer. --- Scheffler, Israel (1991), In Praise of the Cognitive Emotions and Other Essays in the Philosophy of Education (New York: Routledge), pp. 80-96. Turkle, Sherry (2015), Reclaiming Conversation: The Power of Talk in the Digital Age (USA: Penguin) Mark, Gloria (2023), Why Our Attention Spans are Shrinking, https://www.apa.org/news/podcasts/speaking-of-psychology/attention-spans; Attention Span: Finding Focus for a Fulfilling Life (New York: Harper Collins) Haidt, Jonathan (2024), The Anxious Generation: How the Great Rewiring of Childhood Is Causing an Epidemic of Mental Illness (New York: Penguin Random).
José Antonio Perez-Escobar, Deniz Sarikaya	**3E**	Epistemic butlers and critical thinking	AI systems in education can promote epistemic laziness. Asking an AI for information is very different from a Google search: the former yields specifically processed information (dependent on the information that has been fed to the AI and engineering decisions) while a Google search provides different types of information (including conflicting information) from very different sources. Thus, googling information requires critical skills to work with information while AI tools behave as "oracles" in education contexts. This attitude promoted by the use of AI in education is epistemically harmful, even if such AI systems happen to be so advanced that they mostly deliver accurate information. This is because students, as autonomous epistemic agents, need to learn the important skill of working with multiple information signals and critically "distil" or "denoise" them, thereby creating epistemic products and settling for plurally informed positions. We argue that this skill is even more important in the digital era of fake news and disinformation. Even if AI tools can be put to the use of distilling information for us, they de facto incorporate specific values (e.g., political values) and fabricate facts. This can also be read as a challenge for future systems that needs to be addressed. However, as the priority of general AI systems is not education, it would be rather unlikely that "unnecessary burdens" are implemented. Given the differences between educational practice and real-world practices, such adaptations for general AIs may, in some forms, be counterproductive. We have all heard in math class "you will not have a calculator in your pocket, so please do not use one", even if we all have smartphones in real life. To assess the uses of AI in education and critical thinking, we engage with questions of virtue theory, personal development and how AI literacy is connected to forms of a good life and promote active citizenship. We make the case that it is a key-competence to learn to make polished interfaces and ironed out solutions "rougher" by using the right prompts, or less generally, "rougher"-tuned models, providing an epistemically more diverse and challenging user experience. The arguments above are substantiated with the benefits of having multiple teachers of different opinions in education and the psychological notion of "insight", consisting in the appearance of sudden, deep realisations after minor inputs. An overuse of AI in education would diminish the epistemic virtues that they both entail.
Martijn Wiarda, Tristan de Wildt, Neelke Doorn	**3F**	Do missions change values of mission-oriented innovation projects? A Responsible Research and Innovation perspective	'Transformative' mission-oriented innovation policy aims to redirect innovation, and set in motion socio-technical system change. It could be seen as overlapping with responsible research and innovation, in the sense that transformative mission-oriented policy takes an explicitly value-laden position by aiming for change in a societally desirable direction. However, to date, little research has been done on the effectiveness of mission-oriented innovation policy and the evidence of a mission's directional ability is limited. Our presentation aims to examine whether transformative missions redirect the values embedded in mission-oriented innovation projects. To do so, we use the EU Mission 'Restore our Ocean and Waters' and use topic modelling and thematic analyses to identify, conceptualize, and compare latent societal values described in 17 policy documents (i.e., strategic layer), 37 mission-oriented projects, and 809 mission-relevant projects (i.e., operational layer). We subsequently map how these values have changed around the time of the mission launch. The results suggest that the mission's directionality is ambivalent. The mission launch corresponds with an increase of funded mission-relevant and mission-oriented projects, of which mission-oriented projects commonly frame efforts toward mission objectives. However, the results reveal a misalignment between policy and project-level values and show that the prevalence of project-level values remains largely unaffected by the missions. This limited directionality provides a more nuanced understanding of transformative missions. If missions indeed struggle to redirect projects, then this questions their ability to transform socio-technical systems at large.

Irene Niet	**3F**	EU public values governing the twin transition exposed	The European Union (EU) has voiced its support for a twin transition: a joined digital and green transition. Uncertainty exists regarding the norms or public values governing this transition. In the converging energy and IT system, actors are unsure what goals are prioritized, and how to best prepare for the legislations currently under development. Therefore, this presentation aims to expose the EU public values governing the twin transition and critically assess whether these public values embedded in our current institutions (legislations and the competencies of public governing bodies) form a strong fundament for a sustainable, fair, and competitive EU future. This analysis follows a three-step process. First, a literature review of academic articles and technology assessment reports in ethics, energy (transition), digital technologies and twin transition studies public values shows a set of public values which are core to, pressured or supported by the socio-technical developments of the twin transition. Second, To unveil the public values currently governing the twin transition, EU policy documents from 2000 to 2023 are analysed. Third, the two previous findings are compared and critically discussed. What does it mean if the public values determined in academic articles and technology assessment reports as core to, pressured or supported by the socio-technical developments of the twin transition do not overlap with the public values currently embedded in the EU policy documents? In practice, there is some overlap between the two sets of public values. Both academic articles and technology assessment reports and EU policy documents regard sustainability, reliability, affordability and security as important public values in the twin transition. There are also differences. The public value of privacy is only mentioned in EU policy documents in relation to security; not as a separate human right. Academic articles and technology assessment reports also often mentioned balances of power, equity and equality, autonomy and control over technology as four distinct, additionally important public values. In EU policy documents, these values are mentioned much less frequently and often in connection to one another. Balances of power and equity and equality are deemed as similar public values (regularly under the umbrella term of 'fairness'); autonomy and control over technology are grouped as 'control' of technology, whether that is on an EU, national or individual level. Furthermore, the EU policy documents mention public values which were not identified by academic articles and technology assessment reports: efficiency, market integration or unification and services. Efficiency is interchangably connected to several other values: from efficiency in the energy market to ensure greater affordability or reliability, to efficiency for more sustainability, and efficiency leading to more control over technologies. There is not a single definition in the different EU policy documents. Market integration (or unification) and services are often mentioned in connection: a number of EU policy documents deem the integration of the EU energy market, with digital technologies, as a core principle, and this results in opportunities to offer additional (market) services, such as international, accessible, automated bidding systems for the EU electricity market. The differences between public values identified in academic articles and technology assessment reports and EU policy documents give reasons for ethical, philosophical, political and practical discussions. As for the ethical discussion, this discrepancy between public values in practice and policy may indicate a public values change or shift in the public values. However, it is unclear how this discrepancy turns out. The philosophical discussion can centre around the question of who decides on public values, who is the public which is served by them and how values change in general. The political debate revolves around whether this discrepancy means that the existing line of public values, already embedded in institutions, can be followed, albeit with slight adjustments, or if a more radical shift in public values and prioritization is needed for a sustainable, fair, and competitive EU future. Finally, the practice question considers how academics and technology assessment professionals can make their findings about public values and value changes be heard in a way that these can be reflected in new or adjusted legislation.
Inga-Maria Eichentopf, Hans Kasperidus	**4A**	Education for a Sustainable Future - Bridging Knowledge and Action	In today's complex global systems, characterized by interrelated crises such as military conflicts, pandemics and increasing energy prices, human-induced climate change is probably the most critical challenge for the global community. The latest assessment from the Intergovernmental Panel on Climate Change (IPCC) projects that the 1.5°C warming level will likely be crossed in the early 2030s, underlining the urgency of action. [1] However, addressing climate change requires more than technological innovations alone; it necessitates societal and personal transformations. Thus, integrating technical advancements with societal change processes is essential, calling for a comprehensive approach within educational programs to request reflection on past actions and anticipation of future needs. Yet, the teaching practice shows that students tend to see the problems occurring due to climate change more abstract and its solutions in a mostly technical manner. [2] Moreover, surveys reveal students value systems thinking and simulation for their professional qualification, but struggle with understanding accumulations. [3] Research indicates even highly educated adults lack understanding in simple dynamic systems. [4, 5] Many even fail to comprehend the cause-and-effect between greenhouse gas emissions, atmospheric concentration, and global temperature change. [6] To close these gaps, systems thinking and technology assessment, as a process to forecast the consequences of new technological developments, can provide learning with foresight about the societal, environmental, and economic dimensions of a problem. Additionally, the interplay of topics even extends to companies. A survey of Saxonian employers revealed a shift towards valuing generalists and systemic thinkers. To bridge knowledge and action, the integration of educational tools with practical application of simulation models, like the En-Roads simulator, can provide a significant impact on prospective thinking already on the university level. The talk aims to share teaching ideas and discusses results from the dialogue with companies, students, and colleagues researching on systems thinking in education. References IPCC, 2022: Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution ofWorking Group II to the Sixth Assessment Report of the Intergovernmental Panel on ClimateChange [H.-O. Pörtner, D.C. Roberts, M. Tignor, E.S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, B. Rama (eds.)]. Cambridge University Press.Cambridge University Press, Cambridge, UK and New York, NY, USA, 3056 pp.,doi:10.1017/9781009325844. Kioupi, V., Voulvoulis, N., Education for Sustainable Development: A Systemic Framework for Connecting the SDGs to Educational Outcomes. Sustainability 11(21), 1-18 (2019). Kasperidus H.D, Biber P. 2018. Systems thinking and systems dynamics skills inquiry in an international master's program in Sustainable Resource Management. Record of the 36th International Conference of the System Dynamics Society. Reykjavik, Iceland, August 7-0, 2018. Accessed 15.04.2024: https://proceedings.systemdynamics.org/2018/proceed/papers/P2246.pdf Cronin M, Gonzalez C, Sterman JD. 2009. Why don't well-educated adults understand accumulation? A challenge to researchers, educators, and citizens. Organizational Behavior and Human Decision Processes 108(1): 116–130. Sterman, JD 2010. Does formal system dynamics training improve people's understanding of accumulation? System Dynamics Review, 26 (4), 316–334. Sterman JD. 2008. Risk communication on climate: mental models and mass balance. Science: 322: 532–533.

Albrecht Fritzsche, Stan Kranc	**4A**	Science Fiction in Engineering Education: Learnings from the EELISA Initiative	EELISA is an alliance of European Universities with a strong technical profile to foster excellence in engineering education and skills management. It has started a variety of initiatives during the last years, including a science fiction writing contest for students, accompanied by a series of talks. The aim of this submission is to report from this initiative and to develop a sound theoretical basis on which the relationship between science fiction and engineering can further be explored. A promising philosophical approach in this context can be found in literary genre theory of science fiction, which draws strongly on Bloch's philosophy of technology. A key element of this theory is the notion of the "novum", developed by Suvin (1972, 1979), which describes the plot device used in many works of science fiction to distinguish the setting of the story from the world as we know it. The novum brings about a certain degree of alienation, justifying plots that deviate in one way or another from our lifeworld experience. In many respects, this is relatable to Bloch's notion of utopia as a zone of proximal development, familiar enough that basic concepts still apply, but including a certain degree of novelty resulting from potential scientific or technical advancements. The "techno-scientific futures" addressed by various philosophers of technology during the last years can be considered as utopia in Bloch's sense. For engineering education, the novum as a plot device is particularly interesting, because it establishes a link between innovation and utopia. Dramatic storytelling using a novum adds a layer of reflection to engineering that might otherwise be neglected in professional education – or artificially imposed by ethics courses that are added "on top" of conventional curricula, but remain strange to students. Bloch, E. (1986). The Principle of Hope. Transl. N. Plaice, S. Plaice and P. Knight. Cambridge: MIT Press. Bloch, E. (2000). The Spirit of Utopia, Stanford: Stanford University Press. Söffner, H.G. (2024). The Designed Myth Investigations on the structure and effect condition of utopia. Wiesbaden: Springer. Suvin, D. (1972). On the poetics of the science fiction genre. College English, Vol. 34(3), pp. 372-382.Suvin, D., (1979). Metamorphoses of science fiction: on the poetics and history of a literary genre. New Haven/ London : Yale University Press.
Jon Rueda	**4B**	Reproductive autonomy in the age of artificial intelligence	Artificial intelligence (AI) is increasingly being used in reproductive medicine and in various digital applications on sexual and reproductive health. Recently, these developments have sparked various ethical analyses (Afnan et al. 2021; Coghlan et al. 2023; Rolges et al. 2023; Tamir 2023). Not surprisingly, many of the ethical problems of AI—such as its explanability deficits or the existence of biases—are also present in these AI tools in the service of procreative purposes. However, other issues have been less explored. Precisely, this talk aims to address the impacts of AI on reproductive autonomy, an issue that has not yet received the ethical attention it deserves. The structure of the talk will be as follows. First, a review of the most important AI systems in assisted reproduction and digital applications for reproductive health will be provided. More remote or futuristic options will also be discussed, such as the use of AI to support embryo ranking for selective implantation (based on non-therapeutic traits) in case in vitro gametogenesis allows catapulting the (so far) reduced number of embryos available in assisted reproduction processes (Suter 2018). Second, the concept of 'reproductive autonomy' will be clarified, explaining the ways in which AI can impact it, offering a classification of the different types of AI assistance. Next, the ethical importance of protecting and enhancing reproductive autonomy in cases where AI exerts technological mediation will be argued. Finally, the ideas defended will be recapitulated, and some final reflections will be offered. Keywords: Reproductive Autonomy; Ethics of AI; Ethics of Technology; Artificial Intelligence; Reproductive Medicine. References: Afnan, M. A. M., Liu, Y., Conitzer, V., Rudin, C., Mishra, A., Savulescu, J., & Afnan, M. (2021). Interpretable, not black-box, artificial intelligence should be used for embryo selection. Human Reproduction Open, 2021(4), 1–8. https://doi.org/10.1093/hropen/hoab040 Coghlan, S., Gyngell, C., & Vears, D. F. (2023). Ethics of artificial intelligence in prenatal and pediatric genomic medicine. Journal of Community Genetics, 0123456789. https://doi.org/10.1007/s12687-023-00678-4 Rolfes, V., Bittner, U., Gerhards, H., Krüssel, J. S., Fehm, T., Ranisch, R., & Fangerau, H. (2023). Artificial Intelligence in Reproductive Medicine - An Ethical Perspective. Geburtshilfe Und Frauenheilkunde, 83(1), 106–115. https://doi.org/10.1055/a-1866-2792 Suter, S. M. (2018). The tyranny of choice: Reproductive selection in the future. Journal of Law and the Biosciences, 5(2), 262–300. https://doi.org/10.1093/jlb/lsy014 Tamir, S. (2023). Artificial intelligence in human reproduction: charting the ethical debate over AI in IVF. AI and Ethics, 3(3), 947–961. https://doi.org/10.1007/s43681-022-00216-x

Martina Philippi	**4B**	How to address ethical problems in a multi-perspective context: interdisciplinary challenges of XAI	AI is now being used in many areas. Based on pattern recognition and prognoses, it provides recommendations for decisions or even decisions itself. But the black box nature of AI systems is problematic: It is often not possible to adequately understand how the system came to a decision. This can lead to biases in modeling and decision-making that are ethically not desirable. The concept of Explainable AI (XAI) is intended to remedy this situation. It is already used in the development process of AI systems. Saliency maps, for example, show which information was prioritized during the evaluation, making potential biases and underrepresented factors visible. This helps to design AI systems appropriately, for example with regard to the selection of training and test data, the annotation of data or even the modeling of reality. What is new about this concept is that AI systems should also be understandable for decision-making domain experts and affected people who are not experts in the field of AI development, but need a basic understanding of the system's decisions or recommendations. The information about the system that developers understand must therefore be translated for other target groups. This can take the form of explanations, as they work in interpersonal communication. Nevertheless, this approach gives rise to some risks. For example, an explanation could appear plausible because it fulfills our expectations but is not factually appropriate; it could distract our attention from aspects that should actually be important to us, because it takes our explanatory needs into account, but not those of a vulnerable group. In this sense, tacit assumptions can be made in the modelling process that can lead to undesirable (ethically unacceptable) results. This could lead us to make irresponsible decisions, and it could even permanently undermine our trust in AI systems. For in order to make responsible decisions, this trust must not take the form of overtrust, but is dependent on a reflective approach to the technology in question. Vice versa, a loss of trust is possible if disappointment occurs. In order to tailor explanations to target groups and at the same time address the risks mentioned above, interdisciplinary collaboration is essential. Nevertheless, interdisciplinary dialogue faces major challenges. For when stakeholders like developers, highly specialized decision-makers and those affected work together, it is not trivial how priorities are set. What seems obvious to whom depends on what one is aware of, what is particularly important in the respective perspective and also on what is possible, efficient and can be achieved with an acceptable effort. For this reasons, it is important to agree on an informed understanding of what is required for not only practicable, but also responsible solutions. In my contribution, I would like to show how philosophy, through its peculiarity of fundamentally reflecting on presuppositions of different perspectives, can help to bring together the different ways of thinking and priorities in a multi-perspective understanding and therefore to face the challenge of implementing XAI into different practices in an effective and responsible way.
Ted Limbeek	**4C**	Value-based decision-making in asset management of Urban Bridges and Quay Walls (UBQs)	Urban bridges and quay walls (UBQs) in historic Dutch city centres exhibit significant signs of deferred maintenance, surpassing their functional and/or technical lifespan. Municipalities grapple with integrating new insights and promises of technological and institutional innovation into UBQ asset management practices, as well as adapting to evolving value frameworks and societal expectations for the maintenance and renovation of UBQs to be future-oriented. In this, one of the main challenges lies in the lack of clarity regarding the responsiveness of current decision-making processes and their ability to effectively balance existing and forthcoming societal values that are crucial for ensuring the resilience and livability of cities like Amsterdam, The Hague and Zwolle in the future. This paper explores the intricate and dynamic nature of, sometimes conflicting, societal values that are inherent in UBQ asset management. The aim of this exploration is to provide decision-makers with insights to better identify, evaluate, categorise and prioritise patterns of values within the context of UBQ management. These insights will empower decision-makers in dealing with the heterogeneity of values that are considered relevant. This part of the research will be based on a preliminary empirical study of asset management in the municipality of Amsterdam. The initial phase of this research project not only delves into mapping the conceptualisation and institutionalisation of values concerning stakeholders and actors within the socio-technical system of UBQs, but also contemplates the city as a bundle of values in itself. This juxtaposition allows for a more holistic and inclusive account of values, so to do justice to democratic and ethical concerns. This part of the analysis will involve the exploration of philosophical perspectives on the fundamental meaning of 'the city', the role of citizens therein, the evolving functions of UBQs over time, and the values emerging from these considerations. It focuses on what values are historically and theoretically embedded in the assets that are considered in the management of UBQs and, as such, what values are inherently present in the cases of this research project. Emphasising historical and theoretical underpinnings, this phase aims to precede traditional stakeholder and value analyses, bringing engineers and decision-makers closer to foundational understandings of UBQs and how their value can be perceived and managed beyond conventional metrics such as cost and time.

Augustinus Setyo Wibowo, Johanes Narasetu Widyatmanto	**4C**	Navigating through Hydrogen Policymaking by Means of Poiesis	Abstract: In April 2023, France's Hydrogène de France (HDF) signed a memorandum of understanding (MoU) with Indonesia's State Electric Company (PT PLN) for the development of hydrogen technology in Indonesia[1], [2]. This MoU was followed up by signing a Joint Development Study Agreement (JDSA) between both parties in December 2023 during the COP 28 in Dubai. On the national energy policy level, the rush towards hydrogen production exemplifies an aim towards grid decarbonisation. However, different production methods such as hydrogen extracted from coal (bituminous or lignite), or hydrogen from electrolysis which stores excess electricity from solar panels and wind turbines could lead to different carbon profiles of electric grids[3]. Developing hydrogen production in countries with a highly carbonised grid will make the most out of existing fossil-fuelled electricity production, but may not align with the aim of grid decarbonisation such that the latter may just be an add-on to brand high-carbon power plants as 'sustainable'. This work argues that understanding Heideggerian poiesis can facilitate policymakers in considering various grid decarbonisation methods at the local level and scale them up to decarbonise the national grid in a more sustainable way despite the grid's current dependency on fossil fuels. Heideggerian poiesis is an activity to "bring something forth" which is not limited to techne in 'technology' as activities to produce and manipulate what is visible, but furthermore, an activity to unveil physis or nature itself[4]. According to Heidegger, unveiling physis is poiesis in the highest sense due to the emphasis of moving beyond manipulating nature and treating it as mere standing resources[5]. Heideggerian poiesis adds a poetic sense to techne[5] such that the latter could be more than just ways to manipulate and produce for a practical end. Consequently, techne consists of activities to continuously bring forth what has always been there in nature itself with a respectful disposition due to realising that nature is irreducible to productive manipulations. In considering hydrogen production methods, understanding Heideggerian poiesis as activities to respectfully bring forth what has always been there in nature would allow us to reframe the policy question. Beyond considering the most economically viable hydrogen production technologies to pair with existing fossil-fuelled plants, poiesis shifts the focus into thinking of low-carbon technology as respectful dialectics to reveal parts of nature via electricity production. Concretely, this means looking for multiple options of hydrogen production methods and multiple combinations of low-carbon technologies to continuously reduce the carbon output from power plants. We present our work in three steps: 1) critiques towards current views of the implementation of hydrogen production technologies; 2) proposing Heideggerian poiesis as a way to successful grid decarbonisation since it stresses not only on the productive-manipulative capacities of the technology, but also on the importance of respectful disposition towards physis in energy engineering; and 3) policy relevance of Heideggerian poiesis to improve the grid decarbonisation via considering various low-carbon technologies and ways to invite citizens towards environmentally respectful energy consumption behaviour. Keywords: hydrogen production, grid decarbonisation, energy policymaking, respectful disposition, poiesis. References [1] [1] 'Pln Signs Mou With French Hydrogen Company To Promote Net Zero Emissions By 2060 | Portal Kementerian Luar Negeri Republik Indonesia'. Accessed: Feb. 08, 2024. [Online]. Available: https://kemlu.go.id/portal/en/read/4633/berita/pln-signs-mou-with-french-hydrogen-company-to-promote-net-zero-emissions-by-2060 [2] [2] 'HDF Energy va décarboner la chimie en Indonésie', France Hydrogène. Accessed: Feb. 08, 2024. [Online]. Available: https://www.france-hydrogene.org/magazine/hdf-energy-va-decarboner-la-chimie-en-indonesie/ [3] [3] 'Grey, blue, green – why are there so many colours of hydrogen?', World Economic Forum. Accessed: Feb. 27, 2024. [Online]. Available: https://www.weforum.org/agenda/2021/07/clean-energy-green-hydrogen/ [4] [4] R. Terzi, 'Technology and the Ambiguity of Production', in Heidegger and Contemporary Philosophy, vol. 8, C. Di Martino, Ed., in Contributions to Hermeneutics, vol. 8. , Cham: Springer International Publishing, 2021, pp. 35–51. doi: 10.1007/978-3-030-56566-4_3. [5] [5] M. Heidegger, The question concerning technology, and other essays. New York: Garland Pub, 1977.
Rafael Coimbra, Édison Renato	**4D**	Generative AI, Mixed Reality, and the Simulation Hypothesis	Recent technological advancements have challenged the hypotheses put forth by Nick Bostrom in his seminal paper, "Are You Living in a Computer Simulation?" (2003), introducing new paradigms for consideration. Bostrom posits that one of three propositions must be true: (1) the human species is highly likely to go extinct before reaching a "posthuman" stage; (2) posthuman civilizations are unlikely to conduct a significant number of simulations of their evolutionary history or variations thereof; (3) we are almost certainly existing within a computer simulation. Bostrom further argues that unless our current existence is within such a simulation, it is improbable that our descendants will ever engage in ancestor simulations. Despite Bostrom's assertions, the accelerated development of novel technologies, particularly in the realms of generative artificial intelligence and mixed reality, has ignited fresh discussions around the hypothesis of societal disinterest once a supremely advanced technological level is attained. Observations reveal not just a public acceptance (with ChatGPT becoming the fastest adopted application in history) but also a significant technological push by major tech corporations (e.g., Apple Vision Pro). It is premature to ascertain whether humanity will achieve a "posthuman" stage as described by Bostrom. Yet, there already exists a computational infrastructure capable of supporting, to some extent, simulated realities with an impressively high degree of hyperrealistic perception. Concurrently, there is a discernible trend towards embracing a world that is, at the very least, semi-simulated. This opens up new avenues for hypotheses concerning life within computational simulations and its implications. This article aims to explore the current technological landscape and its potential to redefine our understanding of reality. BOSTROM, N. Are you living in a computer simulation?, Philosophical Quarterly, 2003. CHALMERS, D. Reality+: Virtual Worlds and the Problems of Philosophy, W. W. Norton & Company, 2022. KURZWEIL, R. The Singularity is Nearer: When We Merge with AI. Viking, 2024. DESCARTES, R. A discourse on Method, Prabhat Prakashan, 2017. BRAUDRILLARD, J. Simulacra and Simulation, University of Michigan Press, 1995.

Matthew Wragg	**4E**	Constructing bridges using boundary objects: a philosophical account of test, assessment and certification in the construction industry.	Construction product certification tells us something about a product; it informs the user that the product is fit for the intended use as defined by the applicable technical specification (EU CPR305/11, Article 2, §14). The Assessment and Verification of Constancy of Performance system within the Construction Products Regulation provides levels of rigour that are required when demonstrating that the product conforms to the standard. These levels of rigour act as a roadmap for manufacturers to demonstrate the repeatability of behaviour of their standardised products (Chhobra, 2020); a roadmap that changes depending on how the essential characteristics of the product will impact on its fitness for its intended use (EU CPR305/11, Article 2, §4). The web of relevant parties that are involved in this chain is complex. Assessors, testers, manufacturers and third-party organisations require information about products that is specific to their role within the chain but is useable by all. The boundary objects (information used for communication between different groups) that develop (Star and Griesemer, 1989; Star, 2010; Spinardi, Law and Bisby, 2023) through these processes serve to allow the relevant parties to communicate, resulting in a simplistic rudimentary claim: that this product is certified. But, due to the iterative nature of artefact design and testing (Downer, 2007 & 2017), this shared information will also include what knowledge has been gained of the products not just in the test environment, but also from the performance of those previous product design iterations in the "real world". In this talk I shall be defending the thesis that due to the iterative nature of test, the boundary objects that are developed through the route to product certification serve to tether the form and the function of a construction product by using judgement in a way that is just as much from experience of previous products in the "real world" as it is from the test environment. These boundary objects, as useful as they are to industry, may have an added external party due to the function of the artefacts they relate to; that being any user who has the appropriate knowledge to engage with them, not just those in industry. I shall do so in a two stage process. By appealing to the work of Kroes and Meijers (2006) and Downer (2007), I shall demonstrate that the route to certification for construction products answers if the design has met the intention of the designer, as well as if the product has met the standard. This meeting of standard and intention is reliant on judgements that are not born from the existence of a standard, but rather from decisions about what is and is not relevant to testing for the standard. Unlike other boundary objects that serve to manage the central tension between viewpoints and findings within a discipline (Star and Griesemer, 1989), boundary objects within the construction industry have the scope to include those with a less than passing amateur knowledge of the industry. It may serve to help those who live in and rely on the spaces constructed to identify potential failures in performance.
Claudia Eckert, Mark Addis	**4E**	Explanatory Frameworks in Complex Change and Resilience System Modelling	Conventional modelling and simulation have made huge progress in optimising flows for particular conditions. However heterogenous flows across system boundaries continue to pose significant problems for efficient resource allocation especially with respect to long term strategic planning and immediate problems about allocation to address particular resource shortages. Hospital systems which have various patient, staff and equipment flows are an important type of heterogenous flows. The approach taken here to modelling such flows is an engineering change prediction one (Clarkson et al. 2004 and Eckert et al. (2004). This enables margin modelling by producing system models in dependency matrices with different linkage types. Change prediction approaches from engineering design can analyse where these bottlenecks in integrated systems would be so that resources can be deployed flexibility to avoid them and address them when they occur. Current state of the art of margin research can be furthered by identifying margins on multiple levels of system composition. It can usefully be complemented by a category theory based approach which allows representation of variable and constant properties of models under changing conditions, and the identification of flows within models (Breiner et al. 2020 and Foley et al. 2021). There is a difference between an explanatory framework for a model and how algorithms within the model work. Category theory is useful for formalising such explanatory frameworks as it can both structure systems and permit analysis of their applications in a complementary way. References Briener, S., Denno, P. and Subrahmanian, E. (2020). Categories for planning and scheduling. Notices of the AMS 67 (11), 1666-1677. Clarkson J., Simons, C. and Eckert C. (2004). Predicting change propagation in complex design. J Mech Des 126, 788-797. Eckert, C., Clarkson, J. and Zanker, W., (2004). Change and customisation in complex engineering domains. Research in Engineering Design, 15(1), 1-21. Foley, J., Breiner S., Subrahmanian E. and Dusel J. (2021). Operads for complex system design specification, analysis and synthesis. Proc. R. Soc. A 477: 20210099.

Philipp Neudert	**4F**	The Politics and Poetics of Imagination: Scrutinizing the 'Quality' of Imagination	Social scientists increasingly attend to technological visions and imagination, hoping to better understand technological projects and their relation to political power (Jasanoff 2015) or to make imagination a vehicle of desirable sociotechnical change (Schneider et al. 2023). Some writers diagnose a 'crisis of imagination' referring to a growing inability to imagine positive futures (Mulgan 2020, 2022, Pelzer and Veersteg 2019). In light of undoubtedly existing positive future visions, such diagnoses seem to suggest that existing positive future imagination is of insufficient quality. However, the question as to what constitutes 'quality' in relation to what criteria, expectations, or other boundary conditions, has received little explicit attention. To explore this question, this contribution examines two prominent methodological frameworks for producing and assessing futures: Transformative Vision Assessment (Schneider et al. 2023) and the Imaginative Collaboration Framework (Finn and Wylie 2021). It is argued that imagination, as structured by distinct methodological frameworks, can be characterized along the lines of politics and poetics. The former refers to the practical and performative role of imagination for advancing, challenging, or enacting particular political agendas. The latter refers to a method-specific meta-imagination of what good and right imagination should be and do, irrespective of its concrete (actual or ascribed) functions or purposes in specific contexts. The contribution thereby draws on, and contributes to, the broader discussion on the 'politics of method' (Savage and Burrows 2007). More broadly, though, it wants to start a conversation about the politics and poetics of methodologically structured imagination, and how they enable and constrain the creation of responsible futures and, obliquely, a responsible present. Bibliography Finn, Ed, and Ruth Wylie. 2021. "Collaborative Imagination: A Methodological Approach." Futures 132 (September): 102788. https://doi.org/10.1016/j.futures.2021.102788. Jasanoff, Sheila. 2015. "Future Imperfect: Science, Technology, and the Imaginations of Modernity." In Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power, 1–33. University of Chicago Press. Mulgan, Geoff. 2020. "The Imaginary Crisis (and How We Might Quicken Social and Public Imagination)." Demos Helsinki. Mulgan, Geoff. 2022. Another World is Possible. Hurst & Company. Pelzer, Peter, and Wytske Versteeg. 2019. "Imagination for Change: The Post-Fossil City Contest." Futures 108 (April): 12–26. https://doi.org/10.1016/j.futures.2019.01.005. Savage, Mike, and Roger Burrows. 2007. "The Coming Crisis of Empirical Sociology." Sociology 41 (5): 885–99. https://doi.org/10.1177/0038038507080443. Schneider, Christoph, Maximilian Roßmann, Andreas Lösch, and Armin Grunwald. 2023. "Transformative Vision Assessment and 3-D Printing Futures: A New Approach of Technology Assessment to Address Grand Societal Challenges." IEEE Transactions on Engineering Management 70 (3): 1089–98. https://doi.org/10.1109/TEM.2021.3129834.
Maximilian Roßmann	**4F**	The different understandings of hype: over-generalization, over-selling, over-promising, over-resonance, and over-shadowing	In the landscape of technological innovation, hype plays a significant role in shaping perceptions, decisions, and futures. Labeling something as "hype" suggests an element of exaggeration in the communication surrounding technology. However, studies on hype vary greatly in their empirical approaches, value judgments, and suggested interventions. This paper posits that our comprehension and resolution of hypes are fundamentally intertwined with our understanding of communication. It explores how different theories – Signal Theory, Rhetoric, Speech Act Theory, Contaigon Theory, and Discourse Theory – provide perspectives for understanding and addressing technology hype as over-generalization, over-selling, over-promising, over-resonance, and over-shadowing, each shaping the attributed responsibility for hype and guiding potential interventions. Information Theory emphasizes that over-generalizing and generic language can lead to inaccuracies in the information transfer between sender and receiver. This occurs when essential keys to decipher information about limitations are missing, particularly due to limitations of communication channel capacity in title fields, article highlights, and publication formats. Rhetoric examines how the over-selling of claims through rhetorical devices – such as breakthrough metaphors, superlatives, doomsday scenarios, and urgency-evoking storylines – can mislead audiences. In this context, hypers fail to reconcile conflicting communication aims, such as maintaining accuracy versus generating interest, or to put their intended purpose into action. Speech Act Theory suggests that over-promising and accidental promises can result from improper execution or neglect of conventions related to giving commitments. This perspective evaluates hype against the backdrop of legal and cultural norms governing research proposals, contracts, and press releases. Contagion Theory posits that over-resonant buzzwords, crowd behavior, and market trends arise from reinforced mutual observations, affective self-expressive behaviors, and naïve trend projections. In this context, hype critiques deindividuation, highlighting the absence of critical engagement among individuals swept up in collective excitement. Discourse Theory addresses how overshadowing narratives, media, and speaker positions privilege certain voices and claims over others. In this regard, hegemonic narratives are critiqued as hypes when they distract from more pressing issues or impede the ability to speak truth to power. With examples and suggested interventions for each perspective, this presentation provides a comprehensive framework for understanding and addressing hype from multiple angles. It illustrates how these theoretical lenses shape the understanding of hype, incorporate underlying values, attribute responsibility, and propose diverse interventions. By unpacking hype, we can navigate misunderstandings, reflect on our own contributions to technology discourses, and derive meaningful implications for assessing technology futures.

Hans Voordijk	**6A**	AI and user agency in civil engineering practice: a postphenomenological approach	AI and user agency in civil engineering practice: a postphenomenological approach Artificial Intelligence through machine learning becomes increasingly important in civil engineering practice and has been applied, among others, in structural design of building infrastructures, time and cost planning of large projects, and risk quantification. Methods based onmachine learning (ML) use large volumes of stored data and identify patterns or relationships within these datasets through a self-learning process.ML requires 'datafication' of a certain aspect of reality: to datafy a phenomenon is to put it in a quantified format so it can be tabulated and analyzed (Mayer-Schönberger & Cukier, 2013). In essence, the majority of ML technologies describe patterns and real-world phenomenon in a fashion not very comprehensible, intelligible or at the very least rationalizable to human (i.e., black box solutions). Given the ever-increasing accuracy of ML technologies in civil engineering practice, the reliance and dependence of humans on ML-based solutions increase. This can create situations where users of ML technologies perceive this practice from a perspective unbeknownst to them. Decisions become the result of a complex interplay between users and technology (Verbeek, 2008).A major question is if one can speak of a hybrid form of responsibility between users and technology. Does AI take responsibility away from its users and does it has its own agency? Or do designers of these technologies bear responsibility? Concepts from postphenomenology, such as technological and composite intentionality, may clarify this perceived hybrid agency between users, designers and AI in civil engineering practice (Fritz, Brandt, Gimpel, & Bayer, 2020; Redaelli, 2022). By using these concepts, the increasingly close relationship between users, designers and AI in civil engineering practice can be examined. The central idea of postphenomenology is that technologies mediate the relationship between humans and the world they experience. Within civil engineering research, only a few studies use postphenomenological approaches thoroughly. Thus, for researchers in this field, much of applying these approaches remains unrealized. We first elaborate on the technological mediation perspective. Second, the use of ML in AEC practice is described in two case studies. One case study focuses on using ML in increasing the accuracy of cost estimations for new projects using tacit knowledge embedded in data from past projects. The other case focus on ML in predicting and assessing building structural design performance. Subsequently, the mediating role that these ML technologies plays between their users and AEC practice is analyzed. The concept of composite intentionality allows us to better understand the role of AI-based ML technologies. Composite intentionality is a result of a specific human-technology relationship. Cost estimations and design decisions are mediated by ML technologies. These estimations and decisions, however, neither are determined by, nor can be made completely independently of, technology. Composite intentionality clarifies this hybrid form of responsibility between user and ML technology, if understood because of adding technological intentionality to the human intentionality in a hermeneutic relation. Keywords: Civil engineering, artificial intelligence, postphenomenology, agency. References Fritz, A., Brandt, W., Gimpel, H., & Bayer, S. (2020). Moral agency without responsibility? Analysis of three ethical models of human-computer interaction in times of artificial intelligence (AI). De Ethica, 6(1), 3-22. Leiringer, R., & Dainty, A. (2023). Construction management research: a community at a crossroads? In A Research Agenda for Construction Management (pp. 1-20): Edward Elgar Publishing. Mayer-Schönberger, V., & Cukier, K. (2013). Big data: A revolution that will transform how we live, work, and think: Houghton Mifflin Harcourt. Redaelli, R. (2022). Composite Intentionality and Responsibility for an Ethics of Artificial Intelligence. Scenari(17). Verbeek, P.-P. (2008). Cyborg intentionality: Rethinking the phenomenology of human–technology relations. Phenomenology and the Cognitive Sciences, 7(3), 387-395.
Stefan Rinner	**6A**	Large Language Models and Linguistic Understanding: A Modal Argument	In this talk, I will present an argument for the claim that large language models (LLMs) can be said to understand the linguistic expressions they produce. The argument can be compared to modal arguments against materialism in the philosophy of mind, which have been developed, among others, by Saul Kripke and David Chalmers. The basic idea of these arguments is that from the possibility that P & Neg-Q is true, where P is the conjunction of all microphysical truths about the universe and Q is an arbitrary phenomenal truth, we can infer that materialism is false. After all, it is widely accepted that materialism has modal commitments. Similarly, I will argue that the claim that LLMs understand (or do not understand, respectively) the linguistic expressions they produce has modal commitments. Therefore, the argument goes, from the possibility that LLMs understand the linguistic expressions they produce we can infer that LLMs do in fact understand the linguistic expressions they produce. But why should we believe that it is possible that LLMs understand the linguistic expressions they produce in the first place? In order to argue for the possibility that P & Neg-Q, modal arguments use the route via conceivability. Accordingly, if it is conceivable that P & Neg-Q, then it is possible that P & Neg-Q. Roughly speaking, conceivability means that P & Neg-Q expresses a coherent hypothesis: one that cannot be ruled out a priori (even on ideal rational reflection). Hence, modal arguments against materialism have the following structure: (1) It is conceivable that P & Neg-Q. (2) If it is conceivable that P & Neg-Q, then it is possible that P & Neg-Q. (3) If it is possible that P & Neg-Q, then materialism is false. (Therefore) Materialism is false. In this talk, I will present similar arguments for the claim that LLMs, such as ChatGPT (or at least certain further developments of it), understand the linguistic expressions they produce. The arguments will be as follows, where X stands for a certain LLM: (1) It is conceivable that X understands the linguistic expressions it produces. (2) If it is conceivable that X understands the linguistic expressions it produces, then it is possible that X understands the linguistic expressions it produces.

Sabine Thuermel	**6A**	Dancing with Generative AI	Everyone is talking about generative AI, but how can it be used responsibly today? To answer this question, its key characteristics are introduced. Thus, it becomes evident, that we are not dealing with a complicated approach that can be mastered with expert knowledge. Rather, we are confronted with complex systems that require a very specific approach, as will be shown below. Generative AI can learn patterns, e.g. speech patterns, at extremely high speed, depending on access to training data, e.g. texts. It is therefore pre-trained. Depending on the input (usually in natural language), stochastic algorithms generate new digital content that is "as similar as possible" to the input. Due to this non-deterministic approach, every generative AI system exhibits weak emergence: Each time it is used with the same input, it can produce a different output that cannot be predicted, but only emerges during runtime. The first thing to realize here is that generative AI-empowered systems are complex systems. This means that the challenges they pose need to be addressed systematically with no hope of mastering and controlling them once and for all. As Donella Meadows said in her 2008 primer on complex systems in general, the best you can do is to influence them or learn to "dance with them". In the case of generative AI, this means that a multi-stage approach should be used: first, treat generative AI as a tool for specific tasks: experiment in a sandbox. In the next step, use generative AI as an interaction partner once you have agreed on suitable guard rails. Ultimately, strengthened by experience, agree on principles to harmonize AI and humans. Valuable experience can be gained in each step: exemplary are generative agents as realized by Park et al. in 2023: it is demonstrated in a sandbox of 25 agents may display their sociality when organizing a party in virtual “Smallville” all by themselves in a simulacrum of human behavior. However, the expectations of generative agents and their social behavior should remain modest: our living environment, into which humans are born without dependence on training data, offers completely different development possibilities than generative agents possess based on machine token operations and large language models (LLMs). Even today LLM-based agents may be found in a wide variety of application fields. The currently realized scenarios encompass a range of contexts, including a single software agent executing tasks, collaborative and competitive systems of software agents, and software agents assisting humans. If one considers the fact that current AI systems are engineered by humans, Floridi’s dictum that AI is a “divorce (not a marriage) between agency and intelligence” rings true for all current systems. Even these few examples let us be very careful to use generative AI as an interaction partner and when thinking about principles to harmonize AI and humans, we must do everything we can to strengthen our power of judgment, when deciding where and when to use these artifacts.
Udo Pesch	**6B**	The Acceptability of Technological Debates	Discussions on the acceptability of new technologies usually pertain to technologies characterised by something ‘public’, in the sense that they invoke aspects of collectiveness, visibility, or openness. For instance, a technology like solar radiation management does not yet exist, and it may never be realised; still, the very idea of the technology stirs up controversy. ‘Private’ technologies, on the other hand, raise little societal discussion. Their acceptability is relegated to the market domain, in which individual consumers decide about the desirability of products. Smartphones, for instance, happen to have invasive real-life effects. However, they have only become a topic of discussion after their negative consequences become clear, such as the general addictiveness or their effects on school classes or traffic. Also, in the case of technologies that are topics of societal discussion, a bifurcation can be observed in which an understanding of ‘publicness’ plays a crucial role. Either the ‘public’ is seen as a community that is allowed to say something about a specific application within a certain local context, such as the positioning of wind turbines, or the public is seen as a larger homogeneous entity that is asked to evaluate the technology as a whole, for instance by having a voting procedure or a survey, without discussing concrete applications, for instance, nuclear energy. A number of insights can be inferred from these categorisations of technology debates. To start with, societal debates tend to address the effects of technologies instead of targeting the conditions for design and social preconditions. This tendency arises from the idea that technologies are perceived as isolated wholes that are developed outside of society. They are implemented as functional entities that are not subjected to further development. In short, technological applications are not considered elements within interdependent systems that are broader and dynamic. This implies that there is a mismatch between the way public debates are organised and the way technological development takes place in empirical reality. The consequence is that ‘publics’ are only allowed to speak on a limited range of aspects concerning new technology; this explains the rise to societal discontent but also misses out on the opportunity to improve the workings of new technology by being informed by societal values and preferences. To have productive societal discussions on new technology, dominant ideas about the workings of technologies need to be revised, taking the starting point of technologies that develop within an existing sociotechnical setting. In this paper, the conditions for a public debate on technology development will be sketched out so more acceptable debates can be organised.
Yunxuan Miao	**6B**	The (Re)Formation of Collective Values in Technology Development: Saliences and Agenda Setting	Collective values play a crucial role in the development of sociotechnical systems. They not only highlight the focal points where social, financial, and political resources converge but also carry moral significance for technology development—they shape sociotechnical imaginaries through the design of technological systems. In particular, the relationship between values and technology dictates the allocation of resources within these systems. When a value is deemed more important to a particular technology, perspectives related to that value attract greater attention within the technical system, which is essential for value-sensitive and responsible technology design. However, the translation of individual values into collective values through their embedding in institutions and practices remains an area that requires further exploration. Moreover, collective values are not simply the sum of personal values. From a pragmatist perspective, values always depend on the evaluators, and individual focus and evaluation processes can be shaped by underlying social structures and resources. This underscores the need for targeted discussions within specific contexts and frameworks. To deepen the discourse on collective values, this study aims to analyze how collective values are (re)formed in technology development. This study will investigate the mechanisms behind the formation of collective values in technology development by drawing on related concepts that implicitly refer to a collective focal point fostering the creation of collective values, such as Thomas Hughes' concept of reverse salient. Furthermore, I will argue that the (re)formation of collective values can be achieved through the process of agenda setting—the process by which certain issues are prioritized in public discourse, gaining the attention of policymakers and the general public. When specific problems are highlighted and defined, they often bring underlying values to the forefront. These mechanisms vary depending on the nature and manifestation of the problems and their social context. As drivers of attention, salient problems can manifest as technological accidents (functional failures), presumptive anomalies, or crises. These issues serve as starting points for agenda setting. Different types of problems require varied responses from stakeholders, which can lead to differences in the mechanisms of agenda setting. For instance, in the battery industry, technological accidents such as fires might require immediate technical fixes, whereas environmental crises might necessitate comprehensive policy responses and public engagement. Additionally, the effectiveness and direction of collective attention are influenced by institutional and cultural contexts. Institutions such as regulatory bodies, industry standards organizations, and cultural norms shape how collective values are maintained, determining the prioritization of certain issues over others.

Avigail Ferdman	**6C**	AI, Capacity Deskilling and Public Reason	AI creates a serious risk of "moral deskilling": obviating the need for humans to employ skillful moral judgement, by relegating these judgments to machines. Shannon Vallor argues that the risk is that the very ability that humans have, to form moral judgments and cultivate relationships of dependency and care, will be eroded. The philosophical treatment of moral deskilling is primarily focused on the risks of relegating moral practices to technologies: automated weapons technology, new media practices, and social robotics. In this paper I propose to expand the notion of deskilling into the political philosophy realm: the context of 'public reason'. Public reason is a normative commitment to making binding collective decisions, where citizens recognize one another as equal members of the political community, by appealing to reasons that are justifiable or acceptable to all. This paper reflects on the extent to which AI-powered decision-making might affect the epistemic and moral capacities required for publicly reasoning on policy issues, and on the legitimacy and desirability of AI technologies themselves. If indeed AI creates affordances whereby humans no longer need to cultivate and apply moral skills, then there is a risk that they will no longer be able to fully grasp or adequately apply the practice of 'justification', which is central to public reason. There are two possible ways of unpacking the worry of moral deskilling in the context of AI-powered decision-making. One is to explore the human capacities involved in the practice of justification, and examine to what extent AI erodes them. These include epistemic capacities as well as moral attention, empathy, patience, reciprocity, imagination and joint action. Here I explore the ways in which these capacitates are constitutive of the practice of justification. If AI erodes these capacities, then arguably AI could erode the skills necessary for justification, thereby undermining the possibility of public reason itself. Another strategy to unpack the worry of moral deskilling in the context of AI is to apply a particular conception of public reason that speaks directly to the relational aspect of moral deskilling: the 'civic friendship' conception of public reason. According to the civic friendship conception, membership in a pluralistic political community is constituted, in part, by a practice of public reasoning amongst citizens with regard to political rules and institutions. There is value in the joint commitment to making political decisions on public grounds. Yet this value might be eroded when citizens no longer jointly make political decisions, if AI makes collective decisions for them. The aggregation of AI micro-ethical decisions on loans, incarcerations, driving, medical treatments and hiring could lead to the erosion of a joint commitment to making political decisions on public grounds. Here the challenge would be to determine if there is anything inherent in AI large-scale decisions that obviates joint human decision making, which would in turn erode the conditions for practicing civic-friendship public reason. The upshot would be that if we care about public reason, we have reason to limit or regulate certain uses of AI that ultimately lead to moral deskilling in the context of public reasoning.
Maarten Franssen	**6C**	How the philosophy of technology can support the politics of technology	During the past few years it has come to be recognized that, in order to address many of the normative questions involving technology, the approach through ethics needs to be complemented by an approach through politics – as I have argued myself at the fPET 2023 conference. The framework of ethics is centred upon individual human action, but concerning technology, much of what happens, from the design and development of technology, through its implementation, to its regulation and use, takes place at aggregate levels where individual action decides very little. This raises the question how to proceed. Evidently proceeding must involve opening up to 'politics' as a philosophical field, just as ethics of engineering and technology cannot but involve 'ethics'. However, politics and ethics as philosophical disciplines, although on a par to well within the nineteenth century, are currently in very different states, as a result of having been quite differently affected by the critique from empiricist and analytic philosophy in the first half of the twentieth century. Whereas ethics, through a renewed attention to agency and motivation and through the development of metaethics, emerged strengthened from this period, politics became confused concerning the interplay of descriptive and normative elements and fragmented into political philosophy, political theory and political science (see e.g. Plant 1991). In particular, and that is the first part of what I will argue, politics as philosophy kept focussing on the discussion of a few classical concepts – liberty, equality, justice – and ignored issues of agency. But exactly such issues are of the utmost relevance to the politics of technology. Sovereign states and incorporated multinational business companies are agents on the one hand of paramount importance to many forms of decision-making (design, regulation, implementation and use) of technology and accordingly as inescapably involved in any issue in terms of 'what should be done', and on the other hand agents of a supra-individual type that we lack an adequate and convincing account of. Political philosophy is naive and underdeveloped in its treatment of sovereign states and seems to treat quasi-sovereign multinational corporations as of no particular importance at all. One way to come to understand how fundamental these issues are is to look into how Rawls struggled throughout his career with incorporating the existence of countries into his philosophical framework. The second half of my argument is that, fortunately, it is within philosophy of technology itself that resources can be found to help us develop a deeper understanding of states and corporations. I will sketch how the concept of sociotechnical system, originating in the 1950s but rejuvenated from the 2000s on (see e.g. Franssen & Kroes 2009), can provide an entrance to understanding the 'machinery' of society, among which countries or states and corporations figure prominently, and in particular understanding the type of agency that this machinery displays. F. Little, Modern political thought, Blackwell 1991. M. Franssen & P. Kroes, 'Sociotechnical systems', in A companion to the philosophy of technology, Berg Olsen, Pedersen & Hendricks eds., Wiley-Blackwell, 2009.

Malvina Ongaro, Daniele Chiffi, Lorenza Petrini	**6D**	A pragmatist approach to causation in multi-risk research	The nature of causation is a classical topic in both philosophy of science and metaphysics. While most of the discussions have focused on questions concerning the reality of causal relations or the meaning of causal concepts, a pragmatist stance on the philosophy of causation may look at the role that causation plays in the practice of those using it (Price 2001). From this perspective, philosophers of science can ask descriptive questions about which interpretation(s) the scientists actually use (DeVreese 2009) and normative questions about which one(s) they should use in relation to their aims. This approach is particularly relevant in interdisciplinary fields, where methodologies, conceptual frameworks, and ontological commitments must be made compatible at least to the extent required to make the common efforts mutually understandable. Given their importance, assumptions on causation are among those that need to be scrutinised. Aven & Zio (2014) have argued for the necessity of this sort of enterprise in the case of disciplines dealing with various applications of risk, which have often developed differing approaches to the same topic. Anjun & Rocca (2019) take up the challenge and propose that, in the context of risk, causation should be understood in dispositional terms. According to them, this allows the resolution of some methodological problems that the field faces and that are linked to assumptions about causation that follow the Humean tradition. However, in this paper we argue that, at least in contexts of multi-risk, monist accounts like the dispositional one supported by Anjum & Rocca are not adequate from a pragmatist perspective. We argue for our claim by looking at how causation is interpreted in the research on landslides and earthquakes. We identify different approaches both within the single risks and across them, supporting a pragmatist type of causal pluralism. When studying landslides, engineers adopt a propensity account of causation, as they focus on the identification of predisposing factors that determine the area's susceptibility to the triggering of landslides. However, they pair this dispositional perspective with a mechanistic one when they construct computational models that simulate the characteristics of specific landslides. On the other hand, in earthquake engineering the dispositional perspectives is mostly absent, and priority is given to the construction of mechanistic models. These differences notwithstanding, both fields also apply a regularity theory of causes in computing frequentist probabilities for the occurrence of an event in a specific area on the basis of historical data. These different approaches can and should be justified with respect to the aims of the discipline. We argue that this pluralism is both epistemically and practically necessary. With reductive monist approaches, some crucial information about landslides and/or earthquakes would be lost, thus weakening our understanding of the two phenomena. Furthermore, the ultimate aim of multi-risk models is, arguably, to inform decision-making, and the adoption of different approaches permit the identification of the factors on which mitigation strategies should focus in the two cases. References Anjum, R. L., and Rocca, E. (2019). From ideal to real risk: philosophy of causation meets risk analysis. Risk Analysis, 39(3), 729-740. Aven, T., and Zio, E. (2014). Foundational issues in risk assessment and risk management. Risk analysis, 34(7), 1164-1172. De Vreese, L. (2009). Disentangling causal pluralism. In R. Vanderbeeken & B. D'Hooghe (Eds.), Worldviews, Science and Us: Studies of Analytical Metaphysics (pp. 207–223). World Scientific Publishing.
Daria Jadreškić	**6D**	Inductive risk meets engineering risk: What can quality control in engineering teach us about managing values in science? Lessons from studying technical reviewing at CERN	Inductive risk is the risk of being wrong in accepting or rejecting hypotheses, and for at least two decades it has been fueling the philosophical debate on the role that non-epistemic values should play in hypothesis acceptance. Broader notions such as epistemic risk and revived old problems such as demarcation, in this case between legitimate and illegitimate influence of non-epistemic values in science, have been introduced to nuance the debate and do justice to the argument from inductive risk. The argument from inductive risk allows the influence of non-epistemic values in setting evidential standards, especially in the context of applied science such as, for example, in establishing exposure thresholds for chemicals. Such a predicament opens up the need for an appropriate model of value management in science, and different suggestions have been put forward. Inductive risk has not been much discussed in the context of engineering and technology development, mostly because of the recognition of inherently value-laden practices therein, which brings us to the main question of this contribution: Could inductive risk and its focus on error preferences in acceptance/rejection be a useful framework to study engineering and technology development? And vice versa, could mechanisms of quality control ubiquitous in engineering, such as technical review, be useful resources for the debate on value- and error- identification and management in science? Technical review is set up to prevent errors, identify constraints, and harness expertise, among else, on values such as safety, speed, and precision; values that should be measurable, comparable, and tradeable across possible scenarios and managed in a designated social-epistemic format. To draw lessons from technical reviewing, I first show how technical review is different from other forms of peer review in science, namely in not assuming the role of gatekeeping but rather aiming at improving the chosen option under review. In the next step, I look at how this role is practiced in the case of CERN and what this could teach us about responsible research and value management in science, following the framework suggested by Resnik & Elliott 2023. The talk is informed by qualitative interviews with physicists involved in technical review at one of CERN's high-energy physics collaborations during an upgrade of a particle detector. References: Resnik, D. & Elliott, K. 2023. Science, Values, and the New Demarcation Problem. Journal for General Philosophy of Science 54: 259-286.

Emmanuel Caillaud, Océane Salignon	**6E**	The integration of ethics education in French "grandes écoles": an option or a necessity?	The French system of Grandes Ecoles is unique in the world. These highly selective higher education institutions exist alongside traditional universities and are part of the Conférence des Grandes Ecoles. Known for their rigorous academic programmes, French grandes écoles produce graduates who occupy influential positions in industry, government and academia. Their main objective was to train managers and engineers for the French economy. Ethics education plays a crucial role in shaping the moral compass of future engineers and managers. However, the integration of ethics into their curricula remains a subject of debate. We have analysed the curricula of the main "Programme Grande Ecole" courses of the top 10 engineering and management schools according to Le Figaro. If Polytechnique proposes a compulsory lecture by a famous philosopher, Professor M. Foessel, most of these engineering schools propose an optional lecture or conferences on ethics. Some engineering courses may also include a chapter on ethics (e.g. robotics at Isae-Supeaero). The first important finding is the great variation in content: ethics courses vary considerably from one grandes écoles to another. Some institutions offer comprehensive courses, while others integrate ethics into existing subjects. The first limitation of the study is that we focused only on the main programmes, commonly referred to as "Programme Grande Ecole (PGE)". This analysis of PGE programmes is more difficult than expected for two reasons: the decrease in the number of core courses and the increase in the possibility of following internal or external programmes (Masters in Management or mobility with partners). In order to have a more complete analysis, we have also focused on other French "Grandes Ecoles" (not in the top 10) that pay specific attention to ethics and philosophy, such as UTC, INSA de Lyon or ADENCIA, which have a specific department or permanent professors in ethics and philosophy. Finally, we propose to develop a global approach to the development of ethics teaching in French grandes écoles, illustrated by specific programmes in different French grandes écoles: - curriculum improvement - Faculty development - Creation of ethics departments These different elements will be discussed and analysed taking into account the specificity of the French Grandes Ecoles. References Didier, Christelle. 1999. “Engineering Ethics in France: A Historical Perspective.” Technology in Society 21 (4): 471–86. https://doi.org/10.1016/S0160-791X(99)00029-9. Hoven, Jeroen van den, Pieter E. Vermaas, and Ibo van de Poel, eds. 2015. Handbook of Ethics, Values, and Technological Design. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-6970-0. Treviño, Linda Klebe, and Katherine A. Nelson. 2011. Managing Business Ethics: Straight Talk about How to Do It Right. 5th ed. New York: John Wiley. Fanny Verrax. Engineering ethics and post-normal science : A French perspective. Futures, 2017, 91, pp.76-79. hal-04346338
Qin Zhu, Beyza Nur Guler, Andrea Gammon, Rockwell Clancy, Scott Streiner, Ryan Thorpe	**6E**	Exploring How First-Year Chinese Engineering Students Perceive Public Welfare Beliefs and Professional Values	Understanding the experiences of first year engineering students is crucial for engineering educators, staff, and administrators for at least two reasons. First, providing an inclusive and engaging first year experience is instrumental for retaining students in engineering fields, especially students coming from underserved communities. Second, understanding the experiences of first year engineering students provides baseline data for subsequent educational interventions within engineering programs. This paper examines how first-year Chinese engineering students perceive public welfare beliefs and professional values. It is part of a larger project that examines the development of moral reasoning and intuitions during their four years of study in engineering across three cultures: United States, Netherlands, and China. In the data we collected from first year engineering students in China, we particularly examined (1) their perceptions of public welfare beliefs through part of Erin Cech’s well known moral disengagement measure; and (2) their perceptions of professional values through an open-ended question about the three values critical for defining a good engineer. This work adds another dimension to the rationale for studying first-year engineering students, namely, to offer a cross-cultural comparison of the potential differences in experiences and perspectives among these students across various cultures. In particular, it will help us understand whether Chinese engineering students start their first-year programs with the same level of enthusiasm and passion as their counterparts in the United States (reported in our earlier papers and in Cech’s work) and the Netherlands (reported in our earlier papers). This paper will offer valuable insights into how engineering students perceive social responsibility within diverse cultural contexts. These findings will contribute to a more culturally responsive approach to understanding the experiences of engineering students as they transition into the engineering profession. Ultimately, this research will inform the design of intervention programs aimed at training engineers who are globally competent and socially responsible.

Lidia Yatluk	**6F**	How to Make a New Science: Organizational Imaginaries of Decentralized Science	In the recent landscape of science, a significant transformation has unfolded with the emergence of the decentralized science movement and DeSci organizations. These entities present a novel approach to scientific organization, boasting horizontal governance structures and claiming to address previously neglected or challenging research areas by delegating organizational management and knowledge production tasks to algorithms, facilitated by blockchain and AI technologies. In contrast to earlier para-academic endeavors, which sought to address gaps within traditional academic structures (Wardrop & Withers, 2014), these novel organizations signify a departure towards reimagining the organization and generation of scientific knowledge. Thus, my objective was to delve into the futures they envision and construct through their organizational frameworks, as well as to assess homogeneity of the movement. To investigate these visions and current practices, I turned to the concept of the sociotechnical imaginary, originally formulated by Jasanoff and Kim (2009). This framework encapsulates the collectively shared conception of an ideal future, encompassing technoscientific assemblages and sociopolitical order (Kucher & Stigson, 2024). Particularly pertinent to our inquiry is the typology of political imaginaries of blockchain, which encompasses crypto-libertarianism, crypto-commonism, crypto-governmentalism, and crypto-collaborativism (Husain, Franklin & Roep, 2020), along with the sociotechnical imaginary of open science, rooted in Merton's norms of science (1949, 1973). I analyzed organizational documentation from 25 active DeSci organizations. Specifically, I closely examined whitepapers, which are considered primary sources by blockchain researchers due to their comprehensive information on problems, solutions, and technologies (DuPont, 2017). This examination provided insights into the values and norms of each DAO, as compared to proposed blockchain and open science imaginaries. The analysis unveiled a heterogeneous landscape within decentralized science, characterized by various organizational types: crypto-libertarianism (Immortality), crypto-commonism (HairDAO), and open science (JocelynDAO). Additionally, hybrid forms such as commonal-corporate (GenomicDAO) and commonal-open science (CannabisDAO) emerged. Despite diverse orientations, these organizations shared a common aspiration for socio-political autonomy while exhibiting limited critical examination of their technological dependencies, particularly concerning standard protocols and the Ethereum platform. References: DuPont, Q. (2017). Experiments in algorithmic governance: A history and ethnography of “The DAO,” a failed decentralized autonomous organization. In Bitcoin and Beyond. Routledge. Husain, S.O., Franklin, A. & Roep, D. (2020) The political imaginaries of blockchain projects: discerning the expressions of an emerging ecosystem. Sustain Sci15, 379–394. Jasanoff, S., Kim, SH. (2009) Containing the Atom: Sociotechnical Imaginaries and Nuclear Power in the United States and South Korea. Minerva47, 119–146. Kuchler, M., & Stigson, G. M. (2024). Unravelling the ‘collective’ in sociotechnical imaginaries: A literature review. Energy Research & Social Science, 110, 103422. Wardrop, A., & Withers, D. (Eds.). (2014). The Para-Academic Handbook: A Toolkit for Making-Learning-Creating-Acting. Intellect Books. Merton R.K. (1949) Social theory and social structure. New York: The Free Press. Merton R.K. (1973) The sociology of science: Theoretical and empirical investigations / Ed. and with an intro. by N.W. Storer. Chicago and London: The University of Chicago Press.
Sabine Ammon, Nils Neuhaus	**6F**	Metaphors as Conceptual Vehicles for Ethical Vision Design: A case study on the role of metaphors in research and development	The role of technovisions in early research and development can hardly be overstated (see Jasanoff and Kim 2015; see Lösch, Roßmann, and Schneider 2021). We want to utilise this central role of visions on a project level by implementing a method called Ethical Vision Design (EVD). EVD is a facilitated, interdisciplinary process of creating and maintaining a vision together with researchers and other stakeholders to guide development processes in a normative way. Values and goals are made explicit and discussed, resulting in a shared understanding of where the research project should or should not lead, which can then be used to achieve a value change through interventions if necessary (see van de Poel 2022). The presentation will give a concrete example of an EVD process by providing insight into the ongoing biotechnological research project GlobalResist and its development of a biomedical device to test for antibiotic resistance. In showing how our visions concerning development in the area of antibiotics are influenced by the metaphors researchers and medical practitioners use to frame our relationship with bacteria, we will demonstrate how existing visions inscribe themselves into concrete technological development processes through metaphors. If these metaphors are not made explicit, even visions which are not consciously endorsed by the researchers can have an influence on the development process. Since visions and values are often implicit and elusive, they can be hard to integrate into an organised development process. As argued by George Lakoff and Mark Johnson, such abstract phenomena are usually conceptualized through metaphors which tie them to experience, thereby rendering them more tangible (see Lakoff and Johnson 2011). These metaphors constitute an important aspect of scientific as well as philosophical discourse. They thereby influence the development and persistence of technovisions, as seen in ubiquitous metaphors framing the administration of antibiotics as an act of war against the microbial enemy (see Maccaro 2021). Such latent imagery often remains unnoticed, finding its way into our visions without undergoing examination. The link between visions and metaphors has previously been explored in areas like geopolitics (see Yanık 2009) and corporate strategy (see Schenck 2010; Walz 2014). Building on the methodology of systematic metaphor analysis as established by Rudolf Schmitt which aims to locate and analyse salient metaphor fields, we want to include the discussion of metaphors in an ethical vision building process designed for research and development (on the connection between metaphor and moral change, see Eriksen 2020). We will thus showcase our ongoing development of a methodology to address metaphors in EVD. The methodology consists of qualitative interviews, metaphorical discourse analysis and workshops. Within these workshops, the ethicists function as facilitators while also actively participating as members of the research group. It is important to stress that the ethical vision is not imposed by the ethicists. It must instead be understood as the jointly conceived product of a creative and communicative process – only then can it fulfil the desired function of guiding the development process in a normative way. Literature Eriksen, Cecilie. 2020. Moral Change: Dynamics, Structure, and Normativity. Basingstoke: Palgrave Macmillan. Jasanoff, Sheila, and Sang-Hyun Kim. 2015. Dreamscapes of Modernity : Sociotechnical Imaginaries and the Fabrication of Power. Chicago: University of Chicago Press. Lakoff, George, and Mark Johnson. 2011. Metaphors We Live By. 6th edition. Chicago, Ill.: Univ. of Chicago Press. Lösch, Andreas, Maximilian Roßmann, and Christoph Schneider. 2021. ‘Vision Assessment als sozio-epistemische Praxis’. Pp. 337–51 in Technikfolgenabschätzung, edited by S. Böschen, A. Grunwald, B.-J. Krings M. A., and C. Rösch. Nomos Verlagsgesellschaft mbH & Co. KG. Maccaro, Jessica. 2021. ‘Be Mindful of Your Metaphors about Microbes’. mSphere 6(3):e00431-21. doi: 10.1128/mSphere.00431-21. van de Poel, Ibo. 2022. ‘Understanding Value Change’. Prometheus (Saint Lucia, Brisbane, Qld.) 38(1):7–24. doi: 10.13169/prometheus.38.1.0007. Schenck, Klaus. 2010. ‘Sprach-Salto aus der Sackgasse’. Organisationsberatung, Supervision, Coaching 17(3):255–68. doi: 10.1007/s11613-010-0195-8. Schmitt, Rudolf. 2017. Systematische Metaphernanalyse als Methode der qualitativen Sozialforschung. Wiesbaden: Springer Fachmedien Wiesbaden.

Aaron Schultz	**7A**	The Domination of Distraction	At some level, everyone has the power to distract. However, this power is not shared equally in a technological society. A disproportionally small group of agents have the technological means to capture, store, and use behavioral data that can be used to distract billions of people. There is a growing body of literature dedicated to this phenomenon. We can link the imbalance of this power to what Shoshana Zuboff calls surveillance capitalism. As Zuboff puts it, "Surveillance capitalism unilaterally claims human experience as free raw material for translation into behavioral data" (Zuboff, 2020, p. 8). It is the logic that drives the attention economy. The simple yet powerful logic is that behavioral data is a raw material that can be harvested, processed, and sold or used by whoever owns it. At least some of the value of the behavioral data comes from what it promises: control. Data merchants, or those directly involved in collecting and aggregating the data so they can sell it, have financial incentive to (a) attract more users whose data can be harvested and (b) increase the amount of data they can extract from existing users. Incidentally, both aims can be better achieved by utilizing the data they harvest to increase user engagement. Zuboff refers to this as the behavioral value reinvestment cycle. Those who do not mine data themselves but have an interest in using it can purchase packaged data from data merchants. They do so with the hopes that it will help them achieve their own aims by having something that will allow them to better control a targeted group of people. While the primary aim of collecting this data is often to increase profits, it can also be used for moral and political aims. One striking example of behavioral data being leveraged for political aims can be seen in the Cambridge Analytica scandal, which was revealed to the public in 2018. A private consulting firm was able to obtain the data of 87 million users from a researcher who had access to this data and lied to Facebook about how they would use the data. Cambridge Analytica services have been employed throughout the world, in over 100 elections and in over 30 countries (Ghoshal, 2018). When the power to use technology that can attract and command attention is imbalanced and there is no substantial protection against being distracted, those who possess this power can use it largely at their discretion. While there are some restrictions on what companies are allowed to do with the data they collect, the rapidly moving pace of technological advancement makes it difficult for the law to keep up. As the methods of distraction increase in efficacy and in scope of deployment, the imbalance of power will continue to grow, and the arbitrary discretion that it is used with will become ever more concerning. This lays the grounds for a kind of political domination to take place, what we can call the domination of distraction. As a means of protecting ourselves from the forces of the attention economy and the unbridled access a select few have to the technologies of distraction, several scholars have argued in favor of a right to attention. Jasper L. Tran (2016) analyzes the right to attention from a legal perspective and argues that we can understand the right as a property right. Anuj Puri (2020-2021) responds by arguing that the right can be understood as a moral right and focuses on the right to attentional privacy. Bartlomiej Chomanski (2023) argues that the right to attention is a component of a more general right to mental integrity, which is itself derived from the right to bodily integrity. Kaisa Kärki (2022) argues that attention is connected to autonomy and introduces important conceptual distinctions between the autonomy of attention and freedom of attention. Kaisa Kärki and Visa Kurki argue that legal regulations related to attention already exist in legal regulation related to "cognitively challenging tasks and specific attentiveness-related duties," (Kärki and Kurki, 2023, p. 14). If we could come to a consensus on the nature of the right to attention, this would help us understand how we should proceed in order to protect ourselves from the growing imbalance of power. However, existing accounts that defend the right to attention have not yet fully explained the risks we face. In particular, we do not yet have an account of the right to attention that properly understands the moral nature of distraction. In what follows, I will defend the claim that distraction is a unique way to modify someone's behavior in a way that always limits freedom. This makes the act of distraction a morally salient action that can be scrutinized from a moral perspective. While the primary focus of this article is dedicated to understanding the moral features of distraction, it is crucial to keep in mind that the moral evaluation of distraction is most pressing within a technological society deeply intertwined with an attention economy. We are inundated with what I will refer to as distraction technologies, which are technologies whose function partly or wholly depends on capturing the attention of its targets via distraction. Distraction technologies are widely used to capture the attention of billions of people. Once come to treat distraction as a moral action, it will be clear that the most serious violations to our attention can be committed by those with access to these technologies.
Enrico Piergiacomi	**7A**	Moral Technology in Francis Bacon: Ancient Atomism and the Sins of Prometheus	One of the fundamental tenets of ancient atomism is that all our actions should be seen as a means to enhance utility and well-being. Democritus and Epicurus championed this ethical stance, viewing technological progress as morally neutral. According to them, arts and sciences are beneficial if they contribute to a fulfilling life and detrimental if they cause harm or lead to unhappiness. This Democritean-Epicurean perspective significantly influenced the philosophical discourse of the 16th and 17th centuries. In this paper, I will explore how Francis Bacon (1561-1626) interpreted the ethics of technology proposed by his Greek predecessors. Often portrayed as a proponent of selfish utilitarianism, which advocates for the indefinite development of arts and sciences to maximise material gains, he actually held a similar "neutral" conception. He believed that technology must be cultivated efficiently without crossing the line which would make it harmful to nature and human happiness. This perspective is particularly evident in Bacon's allegorical interpretation of the myth of Prometheus in chapter 26 of "De sapientia veterum" (1609). This Titan gifted humanity with fire, which in turn allowed the discoveries of arts and sciences, extensively of technology. Bacon commends this initial phase but criticises humanity's subsequent arrogance regarding its inventiveness ("et perpetuo ad novam industriam et nova inventa extimulantur"). In addition, he argues that it is instead humility that really promotes industry and new discoveries, as it fosters curiosity towards the mysteries of nature, revealing more of its hidden resources for the benefit of humankind. Bacon's view is dialectical in nature. He sees technology as conforming to philosophical utilitarianism only when it remains humble and protects the nature upon which its discoveries ultimately depend.

Nikita Lin	**7A**	Aesthetic Engineering of the Virtual: A Case Study of Jeffrey Shaw's Media Artworks between 1989-1998	What is the virtual? What does virtual reality look like? Since the 1980s, media art pioneer Jeffrey Shaw has demonstrated in his work that "with the mechanisms of the new digital technologies, the artwork can become itself a simulation of reality – an immaterial digital space," and through interactive design "an artwork is each time re-structured and re-created by the activity of its viewers – each person becomes raconteur and auto-biographer of the artwork's many possible forms."[1] From today's perspective, what Shaw had seen as "new digital technologies" are old and outdated technologies, but the conceptual and aesthetic frameworks of simulated virtual space and interactive interfaces that he and his collaborators developed remain relevant and influential in many aspects until the present. In fact, some "futuristic" scenarios the artist constructed using computing and network technologies three decades ago have become part of our daily life. However, this line of artistic thinking and practice – assessing the aesthetic potentials of digital technologies and examining how they can impact our cultural life - has been largely left out in the current ethical and philosophical debates in technology assessment. This paper will present a case study of Jeffrey Shaw's media artworks created between 1989-1998 and investigate into how an aesthetic approach towards engineering, designing, and assessing new digital technologies were developed by the artist and his collaborators. The selected artworks were created while the artist was directing the Institute for Visual Media at the Center for Art and Media Karlsruhe (1989-2003). These pieces were produced under circumstances where the most advanced technical facilities and knowledge were accessible through institutional collaboration. One such example is EVE (Extended Virtual Environment, 1993), a research and development project initiated at the ZKM Karlsruhe in cooperation with the Forschungszentrum Karlsruhe. Built on computer hardware and software available at the time, these so-called "artwares"[2] represent the exhilaratingly new aesthetic expressions enabled by what were then the most cutting-edge computing technologies. Meanwhile, they also exposed technical limitations and failures. In the end, we will return to the question raised in the beginning: What is the virtual? My argument will be that experience of the virtual is determined by technical and engineering limitations, artistic imagination and intuition, knowledge exchange between the artist and the engineers, as well as political and institutional expectations. [1] Shaw, Jeffrey. Virtual Reality: A New Medium for the Artist? In Proceedings of the 2nd Annual Conference on Virtual Reality International: Impacts & Applications. London, 1992. [2] ZKM|Zentrum für Kunst und Medientechnologie Karlsruhe. Hardware Software Artware. Confluence of Art and Technology. Art Practice at the ZKM Institute for Visual Media 1992-1997. Cantz Verlag, 1997.
Joan Llorca Albareda	**7B**	The ethical paradox of automation: the end of human work in the age of AI	The challenge of AI to human work has raised two types of academic debates. First, economists and sociologists are arguing about the feasibility of the future end of human work. They are asking whether the prospects of automation stemming from AI are really possible. We can distinguish two positions: on the one hand, those who defend that AI does not imply a qualitative change with respect to other technologies and that it will produce new types of jobs; on the other hand, those who argue that AI will be able to replace human work in its entirety or to a large extent. Second, some philosophers are exploring the desirability of a world without work. The potential disappearance of human work leads to ethical questions about issues such as the desirability of human work, the value of activities outside of work, or the type of society that would take place under such circumstances. This communication is situated in the second of these debates. My argument will be that, while the end of human work is desirable for human beings, the conditions under which it can take place are morally unacceptable. Full automation of work entails that the technologies that enable it possess a number of properties which, in addition to being economically necessary, are morally relevant. If full automation requires conscious and highly rational AIs, then these entities must be morally considered. That is, we cannot impose on these types of entities to be in charge of performing these activities and freeing human beings from work. The structure of the argument will proceed in the following manner. First, I will offer arguments for the end of human work. On the one hand, I will argue that work is structurally negative in two different senses: in an internal sense, for human autonomy; and in a contextual sense, because of the harms it generates inside and outside the work environment. On the other hand, I will state that work does not possess intrinsic goods, nor is it instrumentally necessary to achieve certain valuable goods: the goods that can be achieved in work can be attained by other, less pernicious, means. Second, I will defend that the critique of work invites us to seek its full automation and not its partial automation. I will adduce three reasons to justify this: i) partial automation does not maintain individual achievement, but rather reduces it; ii) full automation can eliminate the negative aspects of the work ethics; and iii) full automation does not eliminate individual achievement from human life. Third, I will introduce the ethical paradox of automation: while full automation is desirable for human beings, it is morally impermissible. It will be argued that full automation involves enslaving conscious and highly intelligent AIs, which would possess moral status, making full automation of work unacceptable. To this end, I will analyze the arguments of the advocates of the moral acceptability of artificial servants and show that they are insufficient.
Marcell Sebestyén	**7B**	Moral Mirrors: Lessons for AI Ethics from the Shortcomings of Animal Rights	In this presentation, I explore the parallels between the ethical considerations for artificial entities and the historical evolution of animal rights, emphasizing a practical approach to the moral inclusion of machines. Historically, the treatment of non-human animals (hereafter simply called animals) was influenced by a Cartesian view that justified human dominion and dismissed animal suffering. This perspective laid the foundation for modern industrial farming, which even today continues to exploit animals extensively. I will argue that the analogy between animal and AI ethics does not arise from any correspondence between the possible ontological status of machines and animals. On the contrary, in opposition to similar assertions, I contend that the ethical inclusion of AI does not require complex metaphysical arguments about machine consciousness. Instead, it can be justified on the basis of preventing psychological harm to humans. As AI becomes increasingly integrated into social settings, the belief in their sentience and consciousness—whether proven or not—could have a profound psychological impact on humans. The mere possibility that AI could experience emotions coupled with humans' innate tendency for anthropomorphism and the prevalent tendency in robotics to develop highly anthropoid automata could influence human interactions with these systems, potentially leading to psychological distress if these entities are mistreated. This approach highlights the need to prevent potential psychological harm to humans as a result of their interactions with AI. By adopting this practical ethical stance, we can avoid the pitfalls of past ethical failures highlighted by the historical example of animals and proactively integrate moral considerations into the development and deployment of AI technologies. This method not only avoids the ontological debates about AI consciousness but also provides a clear rationale for extending ethical considerations to artificial entities based on their impact on human well-being.

Heike Felzmann	**7B**	Designing for relationship: ethically relevant factors in the design of long-term relational AI	Relational AI is becoming increasingly versatile with the recent improvements in generative AI. This raises interesting ethical questions with regard to ethical considerations and guardrails for the development of relational AI applications, such as long-term AI companions. The experience of relational artificial agents has been investigated extensively, beginning with simple programmes such as ELIZA, to embodied social robots such as PARO or Pepper, and more recently increasingly complex consumer chatbots. Much attention has been paid in research to the identification of design factors that increase user acceptability and positive affect in the interaction, based on a range of affective design elements regarding appearance (for avatars or robots), various expressive elements, communicative features, and interactive behaviour. However, most relational artificial agents have been researched only over short periods of time and are often used in contexts of somewhat guided interaction. Even many robots intended for longer-term use are only user-tested for at most a few months and generally under close guided supervision. There are not many examples of actual long term and independent uses stretching into years: most prominently, these include companion pet robots, such as the robot dog AIBO, and more recently AI companions such as Replika. A peculiar aspects of interaction with relational artificial agents has been their experience as simultaneously 3rd personal, as interactions with a computational object that is clearly understood as an object, and as 2nd personal, as interactions that give rise to experiences of intersubjectivity and attachment to the artificial agent as a meaningful interaction partner. This simultaneity creates potential tensions or disruptions of the relational experience, but also keeps the user aware of the nature of the relational agent as technological system, designed and controlled by humans. One particularly interesting case is when users themselves make conscious decisions determining specific features of their relational artificial agent. Versatile customisation becomes increasingly feasible with AI companions, putting users in the peculiar position of explicitly creating a desired relational experience for which they are both makers and recipients. At the same time, system design decisions and continuous development also affect relational parameters, and may run counter to users' own desires and intentions. In this paper we will explore ethical considerations arising from the longer-term experience and relational impact of active user customisation and its interplay with disruptions associated with system constraints, failures, upkeep and upgrade measures. User experiences with the AI companion system Replika will be used as an illustrative example of the ethical complexity that need to be taken into account by providers of relational AI.
Carl Mitcham	**7C**	Technology Assessment from the Perspective of Political Philosophy	This is an effort to place technology assessment (TA) within the orbit of political philosophy. The thesis is that TA functions as a necessary but problematic element of the liberal nation-state. TA is necessary because of the liberal state dependency on engineering and technology. It is problematic insofar as a technical elite may threaten the fundamental, animating ideal of a liberal regime – that is, liberty or freedom. This problematic drives the development of participatory methods in TA, even as TA can nurture development of its own (sometimes hidden) technical elite. TA practitioners are not unaware of the dynamic, but it is worth deeper political philosophical attention. This claim must nevertheless be qualified: There is no claim that the argument here, which is speculative in the classic theoretical or contemplative sense, has immediate practical benefit for TA or the state. It will not necessarily make TA better. The potential benefit comes from enhancing political philosophical understanding so that we might live more consciously, with eyes wider open, in the world constituted by the TA-practicing liberal state. The argument has three not fully integrated sections. Part one offers brief remarks on TA itself. Part two is a historical-phenomenological sketch of the forces operative in the liberal nation-state. Part three digs in a little more deeply to the ties between the liberal state and engineering. A conclusion engineering and philosophical interests and reiterates the speculative character of the argument.
Hernán Borisonik	**7C**	Exploring the Political Philosophy of Cryptocurrencies	This presentation pursues to offer a nuanced exploration of cryptocurrencies from a political philosophy standpoint, transcending purely economic analyses. Over the past 15 years, cryptocurrencies have garnered significant attention, prompting us to delve into their societal impacts, conceptual roots, and cultural implications. Through categorical frameworks, and comparative analyses, we aim to provide a comprehensive understanding of this burgeoning phenomenon. We scrutinize the historical antecedents and technical underpinnings that precipitated the emergence of cryptocurrencies. This contextualization will try to illuminate the socio-political implications of decentralizing currency and financial systems, challenging traditional power structures and fostering new modes of socio-economic organization. Furthermore, we delineate the distinctions between cryptocurrencies and other forms of digital currency, such as those issued by central banks. This comparative analysis underscores the unique political and philosophical implications of decentralized currencies vis-à-vis state-controlled monetary systems. At the outset, we seek to redefine conventional notions of money, currency, exchange, and the underlying materialities throughout historical epochs. By tracing the evolutionary trajectory of these concepts, we elucidate the transformative nature of cryptocurrencies in challenging established power dynamics. We examine the foundational principles of decentralization, consensus, and autonomy inherent in the crypto ecosystem, reflecting on their political ramifications. Besides, our analysis seeks to extend beyond techno-optimistic narratives to critically evaluate the socio-political implications of widespread cryptocurrency adoption. We interrogate the tensions between individual autonomy and collective responsibility within decentralized networks, highlighting the ethical dimensions of utilizing and promoting cryptocurrencies. Additionally, we explore the moral implications of financial exclusion and the potential for cryptocurrencies to foster greater financial inclusivity, reshaping societal relations and cultural paradigms. In short, this presentation is thought to offer an examination of cryptocurrencies and open a philosophical-political reflection, transcending narrow economistic perspectives. We aim to foster a deeper understanding of the transformative potential and inherent challenges of cryptocurrencies within contemporary society.

Lena Fiedler	**7C**	The ethical problem of gendered robots	Gender equity is a fundamental human right and a necessary foundation for a just world. However, many technical applications and innovations are criticized for not promoting gender equity but harming it. This paper focuses on gendering of robots and argues that this is ethically wrong and a barrier to gender equitable futures. Human-like robots are gendered because their appearance or behaviour evokes stereotypical associations. For example, regarding design, a study has shown that robots with long hair tend to be perceived as women, robots with short hair as men (Eyssel & Hegel, 2012). Although there is an increasing amount of social studies examining gendering of robots (Nomura, 2017), empirical research can only explain the causes and effects of robot gendering. It remains to be clarified if and why gendering of robots is ethically problematic. This paper addresses this question by looking at the problem of representation. A gendered robot is perceived as female and therefore representing the social identity of women. The representation of any social identity – may it be gender, race, class, sexuality, ability or other – is
Paige Benton	**7C**	Democratic AI: Justification for a Broad View of Public Reason	There is an increase in literature discussing how and why AI technology undermines democracy (Coeckelbergh, 2024). One stream of research in AI Ethics is the investigation of the risks of opaque decision-making AI (Fainman. 2019; von Eschenbach, 2021; Chesterman, 2021). The issue of transparency and explainability of AI has led to questions about the ideal method of justification for this technology (Maclure 2021). My talk applies some important and underexplored tools from political philosophy to this question. I will argue that a broad view of public reason as the method of justification for the development and implementation of AI can help align AI technology with democratic practices. Public reason requires justifying decisions to citizens via reasons they can accept, in other words, reasons that do not rely on moral truth claims of comprehensive conceptions of the good. As a result, public reason provides a common standpoint to adjudicate between competing claims of what is just (Quong 2022, Rawls 2005). Democratic societies are characterised by moral pluralism and liberty of conscience, this can lead to moral and political instability when persons' ideals of truth come into conflict. Public reason is an attractive method of justification for democratic societies since it is a form of justification that requires persons to ensure their justification is public insofar as fellow citizens can accept their justification even if they conflict about moral and political truth. The narrow view asserts that public reason should only govern issues related to constitutional essentials, such as using public reason to agree on a constitutional amendment to strengthen the protection of equal rights and eliminate discrimination. In contrast, the broad view contends that the notion of public reason should be employed in all political determinations where citizens wield coercive power over each other (Pariente 2020, 110). On this view, even decisions such as the colour of traffic signs should be publicly justified (ibid., 108). I argue that the public justification for the creation and implementation of AI technology should be justified by a broad view of public reason. AI regulation, if confined to the narrow view of public reason, may lead to many aspects of AI decision-making and implementation not being addressed. This is because many aspects of AI decision-making lie outside of the narrow method of public reason. The narrow view could account for constitutional amendments to strengthen AI regulation. However, the narrow view does not require CEOs of companies to develop publicly justifiable decisions for the creation or implementation of an AI algorithm. This is because the CEO's decisions do not form part of the basic structure in the narrow view. Considering that the use of AI technology is not confined to the public sphere (i.e., basic structure), but instead is integrated into the non-public sphere of society, a broad view of public reason is essential.
Dilek Yargan, Ludger Jansen, Manfred Drack	**7D**	A Conceptual Framework to Understand the Unique Position of Biomimetics	Situated at the crossroads between biology and technology, the areas of biomimetics, biotechnology, and bioengineering – or a more recent approach known as "biological transformation" – develop innovative solutions and products. However, what exactly delineates these research and development fields from each other has not been investigated. In this presentation, we try to state the differences and similarities between biomimetics and its adjacent disciplines and we aim to develop a theoretical foundation for systematising the knowledge generated in biomimetics. The foundation requires that both the epistemology and the ontology of biomimetics are known.As biomimetics seems to have no clearly delineated object of research nor a unified method or objective, our investigation begins with an analysis of biomimetics' epistemic profile, examining its object of research, research approach, methods, objectives, body of knowledge generated, and basic terms. Our general hypothesis is that biomimetics is unique and distinct from other research approaches thanks to these features. We propose a newly developed conceptual framework – comprising function, working principle and construction/design – as beneficial for analysing biomimetics. Using this framework we propose hypotheses on the epistemic profile and test them by means of case studies. For the case studies, we interviewed researchers involved in biomimetic projects, e.g. in developing blast furnace surface structures based on the water transport system of plants, or in developing sensors based on features of seal whiskers to reduce vortex-induced vibrations.The framework is also analysed ontologically, where a focus is on working principles. Working principles are important in engineering design but, so far, have hardly been investigated by philosophers. The case studies are a basis for ontologically investigating function, working principle and construction/design. Additionally, the framework can be formalised to serve as a tool for a shared way of thinking for all members of biomimetic research teams, thereby ameliorating the tension between scientific goals towards gaining insight into biological systems and the engineers'need to build market-viable applications.

Albrecht Fritzsche	**7D**	Who revolves? Around Whom? Reading Günther Anders in the Digital Age	Today, Günther Anders is probably best known for having been Hannah Arendt's first husband and their lifelong written correspondence (). His philosophy of technology receives much less attention. He is not unknown in the fPET community (Miller, 2018), but rarely referenced, and there is evidence that his work was intentionally suppressed in the US (Babich, 2021). Anders was indeed a tough cookie. He spent most of his life as a fierce activist against war and nuclear energy. Traces of this can also be found in his philosophical works. At the same time, they contain a large conceptual repertoire for the analysis of technology and engineering. The aim of our paper is to introduce the most important concepts and discuss their relevance for today's discourse. Anders' main writings on the philosophy of technology have been assembled in two volumes called "The Outdatedness of Man", published in 1956 and 1980. Early onwards, he writes about "Promethean Shame", describing a sort of embarrassment that human beings experience in view of a machine technology that is superior to them. This is further elaborated in the notion of a "Promethean Slope", concerning the inability of human beings to keep up with the possibilities of technology. Connections can be drawn not only to posthumanism or transhumanism (Fuchs, 2017), but also to mundane fixations on machine performance as a reference scale for human performance or humanity as such, e.g. in the idea of a competition with A.I. Unlike other philosophers, Anders does not think that technology expands, complements or fixes the human condition, but rather denounces it as weak. In the second volume of "The Outdatedness of Man", Anders continues with the distinction of three industrial revolutions, but not alongside the criteria that are nowadays commonly used. Anders notices changes in the role distribution of producer and consumer, and the object of production. As the possibilities of technology exceed human abilities to make use of them, us cases need to be produced, i.e. artificial need that call for new technical artefacts. In the next step, human beings themselves produce their own obsolescence, advancing their own replacement and destruction. Similarities exist with shifting experiences of abundance and scarcity in digital technology, as well as reversals of value streams on social media. Approaching engineering on the basis of these concepts can inspire a variety of new accounts of innovation. New actors can be postulated in industrial transformations, for whom moral choice must be described differently. Furthermore, Anders' work can inform a new treatment of digital/ technological sovereignty as a design task. Babich, B. (2021). Günther Anders' Philosophy of Technology: From Phenomenology to Critical Theory. Bloomsbury. Fuchs, C. (2017). Günther Anders' undiscovered critical theory of technology in the age of big data capitalism. tripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society, 15(2), 582-611. Miller, G. (2018). What Ethics Owes Engineering. The Future of Engineering: Philosophical Foundations, Ethical Problems and Application Cases, 229-242.
Natasha Lushetich	**7D**	Energetic Circuits of Non-Binary Machines	In his post-1936 writing, Alan Turing developed the idea of an oracle machine, which, unlike the a-machine, as he called it – a machine that automatically executed all steps of a process until the process was complete – would consult an 'oracle'. Although he did not describe the oracle in any detail, it can be assumed that the oracle would randomly, without resorting to a sequence-locked, procedural logic – configure decisions that would, or might, change the course of the rest of the process (Turing 1939). This means that at the very birth of computation, a different notion of computing, which flirted with randomness, incomputability arising from machinic interoception (Chaitin 2005; Ernst 2018), and unknowability, was considered a possibility. At about the same time Vladimir Lukyanov built a water computer in the Soviet Union. Lukyanov's aim was to understand the spread of thermal mass in metals used in rail track building, particularly their behaviour in extreme weather conditions. He had noticed that water flow was analogous to the distribution of heat and could act as a visual indicator of the thermal process. Lukyanov subsequently constructed a room-size computer out of sheet metal and iron and added glass tubes and tin vessels to monitor the water pressure in the system as water flowed through the glass tubes – which represented stored memory. The water computer produced numerical outputs which determined the answers to the complex differential equations that could be effectively solved with the water computer. Utilising several other examples from early computing, such as Ross Ashby's 1940s robotic homeostat, and W.G. Walter's 1950s machine called speculatrix (a small wheeled automaton), this paper investigates the kinetic and potential energies of machines designed to both deal with uncertainty and unknowability, and uncertainty arising from the random, un-programmed aspects of the machines. A case in point is what is usually referred to as machine speculatrix's self-awareness: the automaton's capacity to 'perceive' itself moving through space by recognising and recording the sources, emissions, frequencies and/or velocities of thermal, light, magnetic and other energies. Discussing situated iterative learning, responsiveness to stimuli, and auto-correction of operative parameters, the paper addresses the bigger question of the relation of particular forms of energy to machinic Umwelts and ecologies. Key aspects of this question are: Can machines that have been designed to deal with uncertainty in an adaptive manner (rather than through yes-no decisions) have a similar level of ecological porosity (understood as openness to the environment) as organic existents? And, how might this change their Umwelt (understood as an entity's perceptual milieu) and Umwelt-based interactions? (von Uexküll 1930).

Daniele Chiffi, Giacomo Zanotti, Viola Schiaffonati	**7E**	A philosophical insight into uncertainty in AI	The notion of risk has been gaining increasing prominence in scientific debates and public discussions on the societal impact of Artificial Intelligence (AI). Most notably, the recently approved AI Act explicitly relies upon a risk-based taxonomy of AI systems to specify the scope of its different levels of regulation. This contribution argues that reflections on AI-related risk can be fruitfully complemented with an analysis of AI-related uncertainty. Although there is no universally accepted definition, in technical contexts risk is often employed to refer to those situations in which the probability describing the likelihood of a certain event are known. However, clear-cut situations of risk, in which we can assign point-like probabilities to the possible outcomes of decisions or states of affairs, are rare. More often, risk involves some form of uncertainty that can be quantified by means of probabilistic intervals, second order probabilities and analogous methods (Hansson, 2018). Such tools can be fruitfully applied in the assessment of AI-related risk, for AI systems are often used in complex contexts influenced by many external factors, in which it may be hard to identify point-like probabilities. However, they might not be enough. By adopting a socio-technical perspective that takes into account both AI systems' technical features and their actual deployment within societies, we emphasize that there are contexts involving the use of AI systems in which we might be unable to meaningfully associate potential outcomes with probabilistic values (Nordström, 2022). Such situations of hardly quantifiable uncertainty often arise when it comes to radically innovative systems, whose large-scale use is substantially unprecedented and for which our estimates can hardly be informed by historical data. Things are even more complex when general-purpose AI systems (GPAIS) are involved. These systems, of which Large Language Models are just the best-known example, are purposely designed to "be adapted to accomplish a range of distinct tasks, including some for which it was not intentionally and specifically trained" (Gutierrez et al., 2023). The problem stems from the fact that different AI systems' uses often involve different possibilities for abuses and malfunctions. This makes it extremely hard to foresee all possible unintended outcomes of GPAIS' deployment, let alone specify their probability, giving rise to contexts of deep uncertainty in which we have to deal with unknown unknowns. AI-related uncertainty represents a significant challenge in our attempts to identify and mitigate AI-related risks, and our epistemological analysis should put us in a better position to cope with different forms of uncertainty. References Gutierrez, C.I., Aguirre, et al. (2023). A Proposal for a Definition of General Purpose Artificial Intelligence Systems. DISO, 2, 36. https://doi.org/10.1007/s44206-023-00068-w Hansson, S.O. (2018). Representing Uncertainty. In S. Hansson and V. Hendricks (Eds.), Introduction to Formal Philosophy. Springer, Cham, 387-400. Nordström, M. (2022). AI under great uncertainty: implications and decision strategies for public policy. AI & Soc, 37, 1703–1714. https://doi.org/10.1007/s00146-021-01263-4
Deniz Sarikaya, José Antonio Perez-Escobar	**7E**	Epistemic diversity in education and AI tools: How we create better tools, by thinking in terms of epistemic virtues	In today's technological landscape, the integration of advanced AI-powered chat bots into education is reshaping the way knowledge is imparted and acquired. However, this talk seeks to illuminate a potential peril that lurks within these adaptations. While the focus is primarily on mathematics, will also briefly consider other areas like biology. This discussion centers around the potential hazards of, among others, standardized video lectures, particularly popular after the Covid-19 pandemic, and their possible role in undermining the principle of pluralism within academia. We argue that the mathematical undergraduate curriculum should not succumb to excessive codification and formalization through the overuse or misuse of online teaching tools. The talk contends that while technology undoubtedly has its place, the rigid structure of such tools risks harming the diversity of thought and pedagogical approaches that has traditionally characterized the academic landscape. In order to underscore the importance of maintaining a diverse array of research practices for fostering mathematical progress, we will present a historical discussion on the role of the Tripos in creating a specific mathematical culture in pre-globalization Cambridge, and the important results that would not have been obtained otherwise. This historical context serves as an example to assess the potential impact of current technological trends on the future of mathematics education and research. After the historical overview, we return to modern times and draw parallelisms between the standardization of education and the concept from economics "winner-takes-it-all": it refers to an economic system where the best performers in a competitive environment outrun the competition and monopolize the rest of the market. This is a term that was studied especially in the context of modern internet-based business models, thus serving our analysis well. Last, we discuss philosophical research stressing how mathematical pluralism is important, from different points of view. This includes not just pure mathematics but also the mathematization of other areas: for instance, there is a variety of alternatives to mathematize biology, each with its own virtues. In doing this, we can better assess what is at stake with an improper homogenization of education, and what the most vulnerable and resilient areas are. This characterization offers us hints to balance technological innovation and the preservation of intellectual richness. One key notion here is productive ambiguity: a certain type of ambiguity is desirable from an epistemic perspective, but the digitalization and homogenization of mathematical notions may undermine it.
Karen Moesker	**7F**	Responsible Innovation Principles in Large Infrastructure Systems: Expanding the 'Leaving Ajar' Approach for Practice	The focus on public acceptance of technology implementation, especially in contentious areas like water recycling for drinking purposes, has led to diverse strategies aimed at acceptance enhancement. Despite significant efforts, the effectiveness of these approaches varies, with criticisms ranging from potential neglect of ethical considerations to superficiality and insufficient consideration of the perspectives of those directly affected. This has left a notable gap in the development of strategies that simultaneously enhance public acceptance and assess the ethical justifiability of these approaches. In response, responsible innovation approaches emerge as potential solutions, assessed through, for example, the 'opening up, 'closing down' and 'leaving ajar' principles. In the context of water shortage in extremely arid regions, where options for enhancing water supply are extremely limited the 'leaving ajar' approach becomes increasingly interesting as it aims to connect efforts to enhance social acceptance while simultaneously considering the responsible innovation paradigm. Yet, the 'leaving ajar' approach currently lacks practical handholds, creating ambiguity in technology development. To fill this gap, this study proposes an assessment framework to connect social acceptance efforts that fall under the 'leaving ajar' approach with ethical acceptability in technology implementation. The working of this framework is illustrated by case studies conducted in the drought-ridden states of Arizona and California.This framework addresses challenges such as, what engaging early means in the context of slowly changing large infrastructture systems, which isa typical requrement of RRI. Moreover, the implications of public engagement process design and their underlying goals are investigated with respect to the feasibility of commonly referred to RRI principles and how these can be applied to such large infrastructure systems.

Beatrice Bonami	**7F**	Principles and Practices of Responsible Research and Innovation towards foundational structures for a technology decolonization methodology in the Global South	The interplay arising between Science and Technology Studies (STS), and Responsible Research and Innovation (RRI) frameworks are under scrutiny by research institutions working on social technology models in the Global South. Such thematic entanglement suggests the need for a decolonial conceptual framework for technologies, asserting that Northern digital industries and corporations often operate within a colonial paradigm, thereby subjecting Southern and non-Western social assemblages (Deleuze & Guattari, 1987) to unjust, inequitable, and unreliable digital experiences and practices (Milan, Beraldo, 2024). Thus, post-coloniality (Mignolo, 2007; Escobar, 2017, Cruz, 2021; Bonami, 2022) reveals pathways to Southern emancipation from contemporary Northern/Western power structures that still dictate interactions within innovation ecosystems. Consequently, research in this intersection demands attention, particularly given (i) the rapid pace of the digital industry, (ii) the resulting imbalance among continents, (iii) the impact on marginalized and vulnerable populations exposed to unfair digital practices, and (iv) the inquiry into the feasibility of decolonizing digital technology admitting a positioned knowledge (Haraway, 1998). This paper explores the possibility of creating a methodology capable of promoting and fostering digital technology decolonization in the development of Artificial Intelligence (AI) (Eke, 2023; Russel, Norvig, 2010) systems in the Global South, drawing insights from mixed methods and case studies conducted in Brazil, Senegal, and Rwanda. The study will adopt a mixed-method approach to understanding a decolonial strategy in developing and appropriating AI systems, with a focus on (i) determining the contextual annotation required for a decolonial method within a given algorithm, (ii) identifying the stage at which this contextual data should be integrated to achieve diverse outcomes, and (iii) examining specific AI models that can facilitate a decolonial AI. Hopefully, these discussions will contribute to conclusions, while assessing whether digital systems can qualitatively be context-driven (Bakhouya et al., 2017) and value-centered rather than merely data-driven. Bakhouya et al., (2017) Towards a context-driven platform using IoT and big data technologies for energy-efficient buildings, CloudTech pp. 1-5. Bonami, B. (2022) Abstraction Clusters to Understand Digital Development: introducing the SETA model. Universität Tübingen Press. Cruz, C.C. (2021) Decolonizing Philosophy of Technology: Learning from Bottom-Up and Top-Down Approaches to Decolonial Technical Design. Philos. Technol. 34, 1847–1881. Deleuze, G., & Guattari, F. L. (1987). A Thousand Plateaus: Capitalism and Schizophrenia (B. Massumi, Trans.). Minneapolis, MN: University of Minnesota Press. Escobar, A. (2017). Designs for the Pluriverse: Radical Interdependence, Autonomy, and the Making of Worlds. Duke University Press. Eke, et al. (eds) (2023) Responsible AI in Africa Challenges and Opportunities - Social and Cultural Studies of Robots and AI. Palgrave McMillan Haraway, D. (1988). Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. Feminist Studies, 14(3), 575–599. Mignolo, W. (2007) El pensamiento decolonial: desprendimiento y apertura: Un manifiesto. Instituto Pensar, p. 25-46. Milan, S., & Beraldo, D. (2024). Data in movement: the social movement society in the age of datafication. Social Movement Studies, 23(3), 265–284. Russell, S. J. 1., Norvig, P., & Davis, E. (2010). Artificial intelligence: a modern approach. 3rd ed. Upper Saddle River, NJ, Prentice Hall.
Karl Dagher, Erik Fisher	**7F**	Ethics as Trojan Horse	Ethical reflection is often taken to be a "micro-foundation" of broader efforts aimed at responsible innovation (Fisher, 2019). Without the ability to infiltrate the techno-economic logic of business settings, however, ethical reflection and responsible innovation may be limited to "window dressing" (Hu et al., 2019). This talk explores a potential route through this impasse. Socio-Technical Integration Research (STIR) has been used in numerous research settings as a tool for measuring and enhancing the capacities of scientists and engineers for embedding social and ethical dimensions into their daily research and development (R&D) work (Conley, 2011; Fisher, 2007; Fisher & Mahajan, 2006; Lukovics & Fisher, 2017; Schuurbiers, 2011; Smolka et al., 2021). Through the analytical lens of Midstream Modulations, the literature shows the effectiveness of STIR in documenting and eliciting changes in thinking and/or behavior of scientists and engineers working in the midstream of R&D, i.e between upstream funding sources and downstream real-world applications of science and technology (Fisher & Schuurbiers, 2013). Although there have been several STIR studies conducted in commercial – or private – R&D settings resulting in accelerated modulations for their participants (Flipse et al., 2013, 2014; Flipse & van de Loo, 2018; Phelps & Fisher, 2011; Zaveri, 2023), few have been able to speed those up to a timeframe that allows the uptake of STIR in private industry as a standardized, scalable tool used for increased corporate social responsibility. To accelerate and standardize STIR, we conducted an experimental study with six participants working on Artificial Intelligence (AI) projects in a multinational technology company. By staggering the STIR sessions with two participants at a time, we modified our approach to the individual STIR sessions in each of the three batches of participants in order to enhance STIR's ability to elicit change. We learned about our participants' challenges in thinking about social and ethical dimensions of their work, and within three weeks, two of our participants exhibited changes in thinking and behavior related to the tasks of their AI projects. Participant 3 (P3) underwent a series of modulations that eventually converged into one deliberate modulation that allowed her to adapt her leadership style from coordinating projects with minimal interference, to advocating for specific approaches to the team's tasks. This change was facilitated by the weekly STIR sessions with P3, during which she reflected on the minuscule decisions she had to take in her work and their potential repercussions on society at large. In another three-week study, Participant 4 (P4) went from following the standard ways of communicating the company-mandated sustainability relevance of his work to his superiors, to a more proactive, informal approach that allows his work higher chances of benefiting society and in turn, his own career progression. In comparison to previous studies (Flipse & van de Loo, 2018), this record time of three weeks confirms previous findings that STIR can bring about meaningful modulations in a private industrial setting. It also suggests that these modulations can occur in a timeframe that could make STIR an effective tool for enhancing the social responsibility of R&D within corporate settings. The rapid modulations address concerns related to time and business value that commonly challenge these corporations' efforts, while simultaneously enhancing participants' abilities to embed social and ethical dimensions into their R&D tasks. The study suggests that STIR can potentially be used as a Trojan horse that allows ethics to infiltrate a company's processes and challenge the primacy of business value from within. References: Conley, S. N. (2011). Engagement Agents in the Making: On the Front Lines of Socio-Technical Integration. Science and Engineering Ethics, 17(4), 715–721. https://doi.org/10.1007/s11948-011-9323-x Fisher, E. (2007). Ethnographic Invention: Probing the Capacity of Laboratory Decisions. NanoEthics, 1(2), 155–165. https://doi.org/10.1007/s11569-007-0016-5 Fisher, E. (2019). Engaging the micro-foundations of responsible innovation: Integration of social sciences and humanities with research and innovation practices. In International Handbook on Responsible Innovation (pp. 194–210). Edward Elgar Publishing. https://www.elgaronline.com/edcollchap/edcoll/9781784718855/9781784718855.00020.xml Fisher, E., & Mahajan, R. (2006). Midstream Modulation of Nanotechnology in an Academic Research Laboratory. In American Society of Mechanical Engineers, Technology and Society Division (Publication) TS. https://doi.org/10.1115/IMECE2006-14790

Joan Llorca Albareda	**8A**	Challenging the agential nature of AI responsibility problems	In this presentation, I will argue that the responsibility gap arising from new AI systems is reducible to the problem of many hands and collective agency. Systematic analysis of the agential dimension of AI will lead me to outline a disjunctive between the two problems. Either we reduce individual responsibility gaps to the many hands, or we abandon the individual dimension and accept the possibility of responsible collective agencies. Moreover, I will adduce that this conclusion reveals an underlying weakness in the various discussions of AI ethics: the habitual lack of attention to the question of the disciplinary boundaries of AI ethics. This absence has made it difficult to identify the specifics of the responsibility gap arising from new AI systems as compared to the responsibility gaps of other applied ethics. Lastly, I will be concerned with outlining these specific aspects. The argumentation will proceed as follows. First, I will develop the general concept of the responsibility gap by attending to both the elements and conditions of responsibility. This will reveal that the responsibility gap arising from the new AI systems depends on the breach of the agential control condition. Subsequently, I will present the three strategies that have been pursued to address the responsibility gap. Some authors have argued that the concept of responsibility is much richer and more dynamic than this interpretation suggests (Tigard, 2020). Another group of authors have argued that the responsibility gap is inconsistent or of little importance (Hindriks and Veluwenkamp 2023). And, lastly, some authors have offered novel conceptual solutions to address the responsibility gap (Nyholm 2018).While each of them has strengths, I will defend that their weaknesses lead us to conduct a systematic analysis of the agential dimension of new AI systems. Second, I will analyze three types of conceptions of AI agency: moral agency with responsibility, moral agency without responsibility, and neither moral agency nor responsibility. I will defend that the first one leads us to understand AI from the concept of responsible collective agency and the last two are reducible to the many hands problem. Finally, I will expose the lack of attention devoted to the disciplinary boundaries of AI ethics and argue that the problem of ethical proliferation (Sætra and Danaher 2022) is at the root of the difficulties that the responsibility gap of new AI systems has generated. After that, I will show the specific aspects of the responsibility gap produced by these systems. From an internal point of view, it creates new objects of responsibility: their large and complex databases, and the criteria used in algorithmic training. From an external point of view, it has greater impact than other technologies: the increase in the number of systems and their ubiquity. References: Hindriks F, Veluwenkamp H (2023) The risks of autonomous machines: from responsibility gaps to control gaps. Synthese 201(1):21. Nyholm S (2018) Attributing agency to automated systems: Reflections on human–robot collaborations and responsibility-loci. Science and Engineering Ethics 24(4):1201-1219. Sætra H S, Danaher J (2022) To Each Technology Its Own Ethics: The Problem of Ethical Proliferation. Philosophy & Technology 35(4):1-26. Tigard D W (2020) There is no techno-responsibility gap. Philosophy & Technology 34(3): 589-607.
Karen Lancaster	**8A**	Capacity, Consent, and AI	Imagine a future where AI provides consent on behalf of individuals who cannot make or express their decisions, encompassing both routine and life-and-death decisions. Despite sounding like a dystopic episode of Black Mirror, I argue that under particular conditions, AI-mediated consent could be an ideal solution. AI could be trained to make decisions closely aligned with the patient's probable choices – doing so impartially (unlike humans). Advances in AI may thus alter not just our practices and policies, but perhaps also our fundamental understanding of consent. Presently, family members provide consent for patients lacking capacity, but the risk lies in the possibility of inadequate understanding of the patient's wishes – or worse, purposely acting against the patient's wishes. Herein, I explore some 4 potential ethical issues which could arise from AI-mediated consent, but I ultimately suggest that AI may offer a viable solution for some patients. The issues I discuss are: Complexity of training AI Extensive training would be imperative for an AI consent-provider to accurately predict and reflect the patient's wishes. For everyday decisions (clothing, food preferences), the AI could monitor and learn the patient's daily choices (just as Alexa can provide personalised playlists). However, for more complex high-stakes decisions such as medical treatments, the AI could perhaps be trained through repeated and highly detailed patient surveys. Lack of insight into the wishes of some patients It is crucial for the AI to be trained on the individual's preferences in order to offer bespoke decisions. Thus, for people who have never had decisional capacity (e.g. those with congenital capacity-limiting conditions) – or who have not trained an AI whilst they were capable – AI should not be relied upon for decision-making. Potential for changing wishes Whilst individuals might change their minds after programming the AI, similar uncertainties exist when family members make decisions based on what their relative used to want. However, unlike human relatives who may forget things or make mistakes, the superior memory and impartiality of AI could mean that it has far greater insight into the patient's nuanced wishes, due to the extensive training process. Concerns about cyber-security AI's susceptibility to hacking, data loss, or leaks are potential problems: it may be easier to hack an AI decision-maker than to coerce the next of kin. However, implementing robust security measures, backups, and monitoring would render the process safer (no such protections exist to safeguard against the possible greed or coercion of family members). AI-mediated consent-giving will not suit everyone, and many people will have next of kin who can make excellent and accurate decisions on their behalf. But for individuals without adequate familial support or connections, exploring the practicality of training AI to act on their behalf during times of diminished capacity could offer a viable option. I therefore suggest that engineers, software designers, medical experts, and legal practitioners – as well as patients whose capacity is at risk of degenerating – ought to practically explore this possibility.

Dane Leigh Gogoshin	**8A**	Taking Charge of AI	Due to assumed problems in the context of autonomous technologies relating to moral responsibility traditionally understood (as backward-looking and concerned with basic desert), e.g. troubling responsibility gaps (Matthias 2004), technology ethicists have begun to focus on forward-looking alternatives (Santoni de Sio and Mecacci 2021; Hedlund 2022; van de Poel and Sand 2023; Gogoshin 2024). The common idea behind these proposals is that the forward-looking functions attached to our practices of holding responsible, e.g. the prevention and reparation of harms, are still possible in these contexts. Roughly, they suggest that the lack of responsibility status of artifically intelligent agents (AIs) can be compensated for by ensuring the involvement of human agents at design and deployment levels, in some case taking responsibility even when they lack control (van de Poel and Sand 2023), or maintaining control in a way that ensures a direct connection to responsibility (Santoni de Sio and Mecacci 2021). Gogoshin (2024) argues that the pressing moral issue about AIs is not a matter of responsibility gaps, which she argues (following the responsibility skeptics) are ubiquitous in traditional, individual contexts. It is rather a matter of the agendas driving the development and deployment of AI. We need not, she suggests, look for responsibility beyond control (pace van de Poel and Sand), since we have, in principle, the control necessary to address this issue. Moreover, we need not pursue the forward-looking benefits associated with our specifically backward-looking responsibility practices – that of enforcing normative boundaries and incentivizing norm-compliance. This is because the right kind of responsibility, if we can call it that at all, is a matter of front-end control and it is intrinsically rewarding. It is what responsibility skeptic, Bruce Waller (2012), calls "take-charge responsibility, which "designates the broader taking of responsibility – including taking charge of one's own plans and projects and life – that must be distinguished from moral responsibility" (Waller 2012: 146). Just as a captain may have role responsibility for a ship, so you may have take-charge responsibility for your projects, your values, your goals, your life' (Ibid: 107). This form of responsibility is a form of control but not one that human agents are automatically endowed with. Developing it depends on many lucky factors. For those who are lucky enough to have it, exercising it is intrinsically rewarding. It is what enables us to take charge of our lives in accordance with our values. In this talk, I am concerned with whether and how this form of responsibility, the responsibility worth wanting, is possible in the context of AI. I begin by canvassing the previously referenced forward-looking proposals and relating them to take-charge responsibility. I then discuss ways in which take-charge responsibility plays out in AI contexts. References Gogoshin, D. L. (2024). A way forward for responsibility in the age of AI. Inquiry, 1–34. https://doi.org/10.1080/0020174X.2024.2312455. Hedlund, Maria. "Distribution of Forward-Looking Responsibility in the EU Process on AI Regulation." Frontiers in Human Dynamics 4 (April 12, 2022). https://doi.org/10.3389/fhumd.2022.703510. Matthias, A. (2004). The responsibility gap: Ascribing responsibility for the actions of learning automata. Ethics and Information Technology, 6(3), 175–183. https://doi.org/10.1007/s10676-004- 3422-1. Poel, Ibo van de, and Martin Sand. "Responsibility beyond Control." In Risk and Responsibility in Context. Routledge, 2023.
Camilla Quaresmini, Eugenia Villa, Valentina Breschi, Viola Schiaffonati, Mara Tanelli	**8B**	Modelling Epistemic Injustice in Innovation Diffusion: A Case Study on Electric Mobility	Algorithmic fairness has become a very important field, as recent works demonstrated the huge impact of biased technologies on a large scale. In response to such problem, the Machine Learning (ML) community proposes a vast number of approaches accounting for fairness and related metrics, with the aim of addressing these issues (e.g., discrimination) in different contexts. Existing approaches generally treat the problem as a mere distributive issue, i.e., reasoning about a fair distribution of resources/opportunities among the population of interest. So far thus the focus has been on ethical elements, such as equity and equality. However, the epistemic side of fairness (namely that related to knowledge) plays an important role, even if it has been sparsely considered in the literature. Indeed, we claim that the extension of existing approaches to epistemic fairness, by capturing peculiar aspects of injustice, is crucial for avoiding discrimination. In particular, in scenarios where information sharing is involved, real social networks may encode bias because trust is not equally distributed across individuals. This is a matter of knowledge, thus concerning the epistemic dimension of the network. The goal of the work is to address algorithmic fairness under an epistemic lens. By relying on the concept of epistemic injustice, as it has been elaborated in social epistemology [1] in terms of the systematic de-authorization of someone as knower, we extend the notion of fairness in an epistemic direction by considering the credibility of the agents in the social environment. Indeed, in a context where agents share information, the message of a speaker can be rejected if the hearer attributes the speaker a lack of credibility (i.e., a credibility deficit) because they do not share the same social features. Admitting the existence of a systematic credibility deficit affecting certain social groups is hence crucial to mitigate this kind of discrimination which is normally neglected. Within a network-based context, in previous works [2] credibility is assumed to be an individual characteristic represented by a parameter that weights the network's nodes (i.e., the individuals). Instead, the objective of the present study is to broaden the analysis by incorporating credibility within relational contexts. Being more realistic, relational credibility determines the weights of the network's edges (i.e., mutual connections), dictating how much each individual is trusted by its neighbors. We discuss this new framework by referring to the specific case study of electric mobility. In particular, we model the spread of the adoption of electric vehicles in a social network as we claim that fairly promoting this diffusion is also an epistemic issue. The introduction of credibility into the diffusion model allows us to investigate the impact of social imitation/discrimination to ultimately optimize resources allocated to spread the adoption of electric vehicles. We show that, in our case study, accounting for the epistemic dimension makes the adoption not only fair but also more efficient, ultimately minimizing policy costs. [1] M. Fricker. Epistemic injustice: Power and the ethics of knowing. Oxford University Press, 2007. [2] C. Quaresmini, E. Villa, V. Breschi, V. Schiaffonati, and M. Tanelli. Qualification and quantification of fairness for sustainable mobility policies. 2023.

Eugen Pissarskoi	**8B**	Ethics of Digital Twins of Online Social Networks	Many caution that online social networks contribute to undesirable social dynamics such as opinion polarization, the spread of fake news, conspiracy theories, discrimination, and large-scale collective outrage. Although these phenomena are well documented in the scientific literature, demonstrating that online social networks have contributed to their emergence has proven elusive. Digital twins of online social networks, TWONs, hold the promise of addressing this problem. A TWON, short for "Twin of Online Social Network," is a computer model that replicates the dynamics of a real online social network (OSN) to such an extent that it can be likened to a "twin" of the actual communication platform. It embodies all pertinent characteristics of the network and its users, drawing upon detailed empirical data to ensure realism. TWONs might be a powerful means to regulate the design of online social networks. They can be used to rigorously quantify social and societal ramifications of platform design decisions, to optimize OSNs based on economic, social, ethical, and epistemic principles, or to conduct rigorous risk assessments of OSNs and informing regulatory efforts aimed at shaping the design of digital communication platforms. However, akin to the technology they aspire to regulate and enhance, TWONs themselves carry the potential for significant adverse impacts on individuals, communities, and societies. A virtual system capable of monitoring and predicting the behavior of a target system, which in turn influences social dynamics, possesses the inherent potential to impact those dynamics itself. With its ability to monitor and predict communicative dynamics within the network, a TWON can be harnessed for diverse purposes. It could be leveraged to pinpoint mechanisms through which misinformation is disseminated more effectively within the OSN. Furthermore, the development of a sufficiently accurate TWON may necessitate use of detailed user data. This information could potentially be exploited against the very individuals the model aims to help safeguard. Additionally, when supplied with user data, a TWON is capable of generating predictions regarding user beliefs, opinions, and behavior. These projections have the potential to encroach upon the fundamental rights of citizens. Thus, if TWONs are successfully developed, they will offer enormous societal benefits, but they will also pose serious societal risks. This raises the question of how a society ought to deal with the development and the use of such a technology. With our paper, we present the results of a first step towards answering this question. We have reconstructed arguments which justify and reject two extreme modes of governance of TWONS: (i) unrestricted availability of TWONs for everybody; (ii) restriction of using TWONs for vested researchers after approving their research project by a governing authority. The reconstruction reveals that both the arguments for and against both modes of governance contain premises which truth-value (plausibility) is currently uncertain for normative and empirical reasons. Based on that, we derive recommendations for how the normative and descriptive uncertainties can be reduced.
Bauke Wielinga	**8B**	Balancing AI Rigidity and the Need for Exceptions: A Virtue-Theoretical Approach	A major concern in AI ethics is that Machine Learning (ML) systems impose determinability on fundamentally indeterminable human lives (Birhane, 2021), thereby imposing a categorization that harms groups that are often already vulnerable. Specifically, the introduction of Algorithmic Decision-Making (ADM) brings with it the risk that rigid ML categories will make it harder to make necessary exceptions to its output. Categories of people who are often overlooked or omitted will likely not be taken into consideration (Star & Bowker, 2007, Benthall & Haynes, 2019), because such systems apply pre-existing rules and categories, and can thus make wrong decisions in unanticipated cases. These worries are tied to epistemic injustice (Fricker, 2007), since they concern an asymmetry in who determines which categories of decisions AI systems use, and what kind of exceptions can be made. Similar concerns come to the fore in recent public debate in the Netherlands. In the “childcare benefits scandal”, government anti-fraud enforcement did not distinguish between intentional and unintentional mistakes, and discriminating fraud detecting algorithms were used (Frankowski et al. 2021). Similarly, Aizenberg et al. (2023) argue that "algorithmic hiring assessments are incompatible with attributes whose meanings are context-dependent and socially constructed". There is a clear challenge to reconcile the rigidity of ADM’s categorization of reality with the many features relevant to decision-making that cannot (easily) be captured by such systems. This paper will use a virtue-theoretical approach (Vallor, 2016) to understand this tension as a conflict between (1) generalism and equal treatment, and (2) particularism and exception-making, to be resolved by taking a middle ground position that skews toward particularism and exception-making. This middle ground position will be formulated by drawing on the work of David Ross (2002), whose moderate moral particularism sees general moral principles as providing only pro tanto moral duties. This account will contribute to epistemic justice by emphasizing respect for the knowledge and perspectives of individuals over homogenous rule-application. Aizenberg, E., Dennis, M. J., & van den Hoven, J. (2023). Examining the assumptions of AI hiring assessments and their impact on job seekers’ autonomy over self-representation. AI & SOCIETY. https://doi.org/10.1007/s00146-023-01783-1 Birhane, A. (2021). The Impossibility of Automating Ambiguity. Artificial Life, 27(1), 44–61. https://doi.org/10.1162/artl_a_00336 Star, S. L., & Bowker, G. C. (2007). Enacting silence: Residual categories as a challenge for ethics, information systems, and communication. Ethics and Information Technology, 9(4), 273–280. https://doi.org/10.1007/s10676-007-9141-7 Benthall, S., & Haynes, B. D. (2019). Racial categories in machine learning. Proceedings of the Conference on Fairness, Accountability, and Transparency, 289–298. https://doi.org/10.1145/3287560.3287575 Frankowski, A., den Uijl, H., Hendrikx, W., Frissen, P., & Huiting, M. (2021). Tussen staat en menselijke maat. Nederlandse School voor Openbaar Bestuur. https://www.nsob.nl/denktank/overzicht-van-publicaties/tussen-staat-en-menselijke-maat Fricker, M. (2007). Epistemic Injustice: Power and the Ethics of Knowing. Oxford University Press. Ross, W. D. (2002). The Right and the Good. Clarendon Press. Vallor, S. (2016). Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting. Oxford University Press.

Diane Michelfelder, Sharon Jones	**8C**	Sustainable Communities and the Challenge of Caring for Future Generations	Recently, the US National Society of Professional Engineers (NSPE) (2023) underscored that an engineer is obligated to "be engaged in creating, maintaining, and renewing sustainable communities." NSPE (2023) stresses that the reason for this obligation is because "sustainable communities are critical to public health, safety, and welfare." By tying sustainability with public health, safety, and welfare, NSPE is presently an outlier among other US engineering societies (forthcoming). How might a mandate for professional engineers to actively promote sustainable communities, which requires that the engineer design for intergenerational equity, become more widespread, and what ethical perspective might work well to support it? The overall aim of this paper is to explore these questions. Put more specifically, this paper looks to make a contribution, from the angle of engineering ethics, to the just-now-budding conversation (see Makoff & Read 2022; Randall 2019) on intergenerational equity and the ethics of care. When it comes to inspiring engineers to be more attentive in their work to intragenerational issues such as environmental justice, the ethics of care, we maintain, holds promise (forthcoming) though we acknowledge that more work is also needed here. But can the ethics of care also hold promise when it comes to making a difference with regard to intergenerational equity? This question is a tricky one. When it comes to meeting intergenerational needs, the ethics of care appears to run up against its limits of usability for engineers. Certainly in the eyes of care ethicists such as Joan Tronto (1993), far from being a unidirectional activity, caring involves responsiveness—taking feedback from the one who is cared-for and making modifications if necessary to better meet the needs of the other. If the ethics of care holds promise for sustainable design, one would have to either a) turn to a model other than that offered by Tronto, which is the direction taken by Randell (2019), or b) modify Tronto's model so that it fits. Van Grunsven et al. (2023) take the second approach in proposing to shift engineering ethics education toward an ethics of care that emphasizes forms of care that recognize that engineers rarely have direct contact with those who will be using the products of their designs. In other words, responsiveness is already lacking. Taking Van Grunsven's approach as a starting point, we examine how intergenerational equity could be bolstered by a Tronto-inspired ethics of care. References: Anonymized for purposes of review. (Forthcoming). Makoff, R. & Read, R. (2022). For a Care-Based Intergenerational Ethic. In S. M. Gardiner, The Oxford Handbook of Intergenerational Ethics. (online edn, Oxford Academic, 10 March 2021). https://doi.org/10.1093/oxfordhb/9780190881931.013.4 National Society of Professional Engineers (NSPE). (2023). NSPE Position Statement No. 05-100-Sustainability. https://www.nspe.org/sites/default/files/sites/default/files/resources/PSdownloadables/Sustainability.pdf Randall, T. (2019). Care Ethics and Obligations to Future Generations. Hypatia, 34(3), 527-545. Tronto, J. (1993). Moral Boundaries: A Political Argument for an Ethic of Care. New York: Routledge. Van Grunsven, J.B., Marin L., Stone T., Doorn, N., and Roeser, S. (2023). How Engineers Can Care From a Distance: Promoting Moral Sensitivity in Engineering Ethics Education. In G. Miller, H.M. Jeronimo, and Q. Zhu,, Thinking Through Science and Technology: Philosophy, Religion, and Politics in an Engineered World. Rowman & Littlefield International, 141-163.
Elliott Woodhouse	**8C**	Environmental Ethics and Philosophy of Technology – The Problem of Artificiality and the Acceptability of Geoengineering in Climate Strategy.	Geoengineering is a somewhat contested term used to refer to two broad kinds of responses to the problem of anthropogenic climate change – carbon capture and solar radiation management. While these techniques work in quite different ways, one uniting factor is that both involve making intentional, engineered, changes to the way core biogeochemical climate features operate. Indeed, 'intentions' are given a key role in the definition of geoengineering interventions, such as the much repeated 'Oxford' definition: "Geoengineering is the deliberate, large-scale intervention in the Earth's natural systems to address climate change."[1] When members of the public are introduced to the idea of geoengineering, they often respond by expressing a worry about the potential of these kinds of intervention to 'mess with' or 'tamper with' nature.[2] A plausible way to interpret this worry is that it is a concern with the human relationship to nature – that geoengineering, if deployed, would go beyond the legitimate remit of human influence in nature. In other words, the worry expressed is a kind of lay environmental ethics. Therefore, if we want to understand these public concerns with these climate technologies, or to include design considerations to ameliorate these worries, we should try to understand environmental ethics responses to geoengineering. A productive area of technology/environmental ethics collaboration has been defining and understanding the moral significance of the natural/artificial boundary. In particular, environmental philosophers such as Eric Katz and Keekok Lee have placed a great deal of moral importance on maintaining the distinction between the natural and the artificial. Both authors argue that the inherent value of nature is a factor of its independence from humans and therefore is lost when we make intentional alterations to natural entities and systems. These theories offer the grounds for a critique of geoengineering – that the harm it causes is of making the natural world increasingly unnatural, engineered, or artificial. This paper takes an applied ethics approach, going beyond simply evaluating these success arguments, but thinking about what kinds of design considerations might be imposed on geoengineering development to best ameliorate these concerns for those who are convinced by them. It considers whether 'nature-based solutions' technological framings of geoengineering are an effective way of doing this. [1] Oxford Martin Geoengineering Program (2018) What is Geoengineering? Why Consider It? https://www.geoengineering.ox.ac.uk/www.geoengineering.ox.ac.uk/what-is-geoengineering/what-is-geoengineering/. Note that this is a fairly standard definition, and the focus on intentionality has been present from very early on. [2]For instance: Cox, E., Spence, E., Pidgeon, N., (2020) Public Perceptions of Carbon Dioxide Removal in the United States and the United Kingdom. Nature Climate Change. Vol. 10. pp.744–749. Corner A.J., Parkhill, K., Pidgeon, N.., Vaughan, N.E., (2013) Messing with Nature? Exploring Public Perceptions of Geoengineering in the UK. Global Environmental Change. Vol. 23. Is. 5. pp.938-947. Wibeck, V., Hansson, A., Anshelm, J., (2015) Questioning the Technological Fix to Climate Change - Lay Sense-Making of Geoengineering in Sweden. Energy Research Social Science. Vol. 7. pp.23–30.

Elisabeth Does, Anna Rifat Klassen	**8C**	Moralizations in the debate about genetic engineering in agriculture	The research project we would like to present at fPET is dedicated to forms, causes, and effects of moralizations in science communication. An interdisciplinary team of researchers from linguistics, communication studies and philosophy investigates those three aspects of moralizations in public debates on food security, artificial intelligence and energy security. At fPET we would like to present the methodology and first results of a case study on genetic engineering in agriculture we are currently working on in the philosophy team. In this study, we first identify the relevant stakeholders of the debate that is a part of the realm of food security, like farmers' federations or a government nature conservation agency. We compile a body of published texts that represent the stakeholders' perspectives and arguments regarding genetic engineering in agriculture. Based on preliminary work of the linguistics team, we establish a systematic analysis of linguistic expressions with normative significance in this particular debate. Our aim is to construct a descriptive taxonomy that delineates various forms of moralization. These empirically based results are analysed in the light of with existing scholarship in moral and political philosophy as well as philosophy of science. On the basis of the taxonomy we conduct a normative analysis that is founded in philosophical accounts on moralizations differentiate between legitimate and illegitimate forms of moralization. Thereby, we respond to the desideratum in philosophy to research specific issues where moralizations occur. The result of our analysis determines criteria for the identification of legitimate and illegitimate forms of moralizations in debates about moral problems in science and technology. The results of the other disciplinary teams combined with the criteria for identifying legitimate and illegitimate forms of moralizations will constitute the foundation for guidelines for effective and responsible science communication for practitioners in that field, for example journalists or scientists.
Eswaran Subrahmanian, Albrecht Fritzsche	**8D**	Approaching Design Theory from the Perspective of Indian Jaina Logic	Formal symbolic logic plays an important role in contemporary design theory. What has received little attention are alternate views of logic in Eastern philosophy, such as the Syadvada – or assertion of possibilities – in Indian Jain philosophy. The Syadvada does not only distinguish between true and false, but also allows for other logical states, expressed in a seven-fold predication (Sarma, 1994). Syadvada does not subscribe to the Aristotelian principle of the excluded third (Schang, 2010). Interpretations of the Syadvada as a seven-valued logic must be handled with care, as its philosophical foundations radically differ from Western thinking (Jash, 2023). The Syadvada allows for incomplete claims or different standpoints regarding truth to coexist (Ganeri, 2002) for design theory, this offers interesting possibilities to distinguish not only between the known und the unknown (Le Masson and Weil, 2013), but also between paths pursued by design and paths not pursued, or conflicts addressed and conflicts ignored. Applications of Western propositional logic require an account of a problem that is free of contradictions. Conventionally, this account relies on binary distinctions between true and false, which are achieved by a suitable framing of a problem in such a way that indeterminacy is excluded. Design propositions must therefore always be considered together with a subtext concerning the choices that are made regarding these exclusions in framing the problem. We discuss the Syadvada as a means to make this subtext visible and accessible for further reflection. Using practical examples, we explore the meaning of the seven alternatives in practice. Overall, the Syadvada can be considered as a way to address design propositions in reference to the point of view taken by the designer in phrasing them. Moving progressively through the seven alternatives is relatable to a process of relaxation in the treatment of problems. It allows us to explore design as a discourse, in which designers differentiate between what is known, what is unknown, what is an inconsistency or paradox, and what is not accessible in these terms (indescribable). Jaina logic suggests that these boundaries are fungible. It holds different options for designers to act, which are easily obscured by the tunnel vision of Western propositional logic. Ding, R. (2020). Taiji Logic: Chinese Wisdom for Project Governance. Singapore: Springer. Fang, T. (2012). Yin Yang: A new perspective on culture. Management and organization Review, 8(1), 25-50. Ganeri, J. (2002): Jaina Logic and the Philosophical Basis of Pluralism. History and Philosophy of Logic 23(4) 267-281 Jash, A. (2023). Many-valued logic and syādvāda. Jain Journal61(1), 35-46. Le Masson, P., & Weil, B. (2013). Design theories as languages of the unknown: insights from the German roots of systematic design (1840–1960). Research in Engineering Design, 24, 105-126. Sarma, V. V. S. (1994). A survey of Indian logic from the point of view of computer science. Sadhana, 19, 971-983. Schang, F. (2011). Two Indian dialectical logics: saptabha˙ng¯ı and catus.kot.i", Journal of the Indian Council of Philosophical Research 27, 45-75.8
Alexander Herwix	**8D**	A Paradigmatic Framework for Responsible Design Science Research	The rapid development of life-changing information technologies (IT) poses a significant societal challenge: finding ways to ensure a form of responsible innovation (Stilgoe et al., 2013) that can harness the potentially significant benefits of emerging technologies while mitigating the risks of harmful or even catastrophic outcomes (e.g., Center for AI Safety, 2023; Jirotka et al., 2017). However, our understanding of how to implement this imperative is still in its infancy, especially in the face of accelerating technological advances, such as in the field of artificial intelligence (e.g., Blok, 2023; Consilience Papers, 2022). In an effort to contribute to the development of a better understanding of the challenges involved in the practical implementation of responsible innovation, this paper draws on the knowledge and experiences of 41 design science researchers (i.e., researchers interested in the development of innovative IT artifacts that address real-world problems; Hevner et al., 2004) from the field of Information Systems (IS), as well as relevant academic literature, to develop a paradigmatic framework for responsible design science research. Specifically, the framework proposes four interrelated, paradigmatic challenges that responsible innovation efforts must address: Ontological challenge: Making warranted boundary judgments, Epistemological challenge: Using, producing, and growing a body of useful knowledge, Axiological challenge: Managing conflicting values and positions, and Methodological challenge: Developing, selecting, and integrating a portfolio of practices. The paradigmatic framework was developed through a multi-grounded (Goldkuhl & Cronholm, 2010) research process involving empirical, theoretical, and internal grounding of the emerging framework. Its main contribution is that it can provide a common ground for a wide range of researchers and practitioners that are interested in addressing real-world problems through responsible innovation. In this sense, it aims to contribute to a more systematic and reflexive discourse on how to successfully implement responsible innovation in research and practice.References Blok, V. (Ed.). (2023). Putting Responsible Research and Innovation into Practice: A Multi-Stakeholder Approach (Vol. 40). Springer International Publishing. https://doi.org/10.1007/978-3-031-14710-4 Center for AI Safety. (2023). Statement on AI Risk. Center for AI Safety. https://www.safe.ai/statement-on-ai-risk Consilience Papers. (2022, June 26). Technology is Not Values Neutral: Ending the Reign of Nihilistic Design. The Consilience Project. https://consilienceproject.org/technology-is-not-values-neutral/ Goldkuhl, G., & Cronholm, S. (2010). Adding Theoretical Grounding to Grounded Theory: Toward Multi-Grounded Theory. International Journal of Qualitative Methods, 9(2), 187–205. https://doi.org/10.1177/160940691000900205 Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science In Information Systems Research. MIS Quarterly, 28(1), 75–105. Jirotka, M., Grimpe, B., Stahl, B. C., Eden, G., & Hartswood, M. (2017). Responsible research and innovation in the digital age. Communications of the ACM, 60(5), 62–68. https://doi.org/10.1145/3064940 Stilgoe, J., Owen, R., & Macnaghten, P. (2013). Developing a framework for responsible innovation. Research Policy, 42(9), 1568–1580. https://doi.org/10.1016/j.respol.2013.05.008

Nico Formanek	**8D**	Scaling things up! The philosophy of technology at scale	Scaling technologies from proof-of-concept to industrialized production is an essential part of engineering. Rather remarkably scaling processes have neither been discussed much in philosophy of engineering nor in philosophy of technology. This is even more astonishing as there are lots of philosophical questions that can be asked about scaling. For example: What makes a technology "scalable"? What kind of knowledge is gained in scaling technologies? What can we learn from failures of scaling? Should we scale? Is there a fixed point where maintenance requirements balance scaling gains? And perhaps the most important question is: Can we know any of this before we attempt to scale? One aim of this talk is to give a first exposition of questions philosophy of technology could ask about scaling. Answers will at best be tentative as we are charting unexplored territory. But we will try to learn some lessons from the scaling of computer technology. The predicted error rates of the first vacuum computers e.g. ENIAC where so high that many engineers believed that they never could be scaled up. Von Neumann developed a theoretical argument against these skeptics which led kick-started field of fault-tolerant computing and error-correcting codes. In the end his argument was not needed because engineering advances suppressed error rates beyond imagination. But it lives forth in the current debate about the practical possibility of quantum computers. Here we do have proof-of-concept but the engineering challenges are manifold. One obvious lesson we can draw from this case is that practical success trumps theoretical considerations. But if practical success is lacking, theory might be the only thing guiding scaling attempts.
Michael Funk, Albrecht Fritzsche	**8E**	Engineering Digital Sovereignty? – A new Deal on The Tacit Source	The concept of tacit knowing by Michael Polanyi is well known in the epistemology of engineering and technology. Polanyi´s claim that “we can know more than we are able tell” became a classical topos within the strong tradition of investigating non-propositional knowledge. For instance, Don Ihde reflects the role of bodily knowledge and its material hermeneutics in technological lifeworlds. Eugene S. Ferguson and Walther G. Vincenti emphasized the role of imagination and the inner eye in engineering. The primacy of propositional knowledge has been overcome such as the reductionism of treating engineering as mere applied natural science and/or mathematics. Additionally, influenced by historical and sociological investigations on engineering practice, the philosophical debates on engineering knowledge became more divers. After all, today tacit knowing seems to be not an innovative topic any more. However, in times of AI we argue for a new deal on the tacit dimension, especially since it serves as a source of engineering digital sovereignty. Tacit knowing has so far mostly been treated as a residual category, filling the gaps in our accounts of actors and their agency. What we would like to do is reverse the perspective in such a way that they are not just the 'loose ends' in our models anymore, but rather the 'loose beginnings', informed by daily practice, which is now radically changing. On the one hand, non-propositional knowledge has been often investigated as fundamental of engineering knowledge in epistemology. On the other hand, in political philosophy for instance privacy is discussed as central societal norm in times of omnipresent AI Systems and algorithmic data processing. We want bring both debates together and show that tacit knowing – especially the diversity of skills, competence etc. – is an essential critical infrastructure. In a world permeated by digital technology, where algorithms support decision-making based on dispersed, intransparent and continuously evolving datasets, we want to raise the question about the design of responsible futures at the intersection of engineering knowledge and digital sovereignty. The tacit basis of engineering has always ben a philosophical topic, but is now increasingly becoming a matter of design, as data access and processing patterns are purposively constructed and negotiated in technical, economic, and political discourses. Therefore, we emphasize the manifold role of tacit knowing as non-propositional, bodily, practical basis of both technical expertise and ordinary lifeworld expertise in democratic societies; as specific object of digitalization, data analysis, machine learning and therefore as specific source of technical functions that ICT engineers design (e.g. profiling of human behaviour, bias in training data etc.) as a fragile critical infrastructure that should gain more attention in the design of responsible ICT with respect to resilient and sovereign, democratic societies (e.g. the diversity of ordinary life knowledge as vital fundamental of democratic processes). At the intersection of epistemology and political philosophy we argue for a new deal on the tacit source.

Kasper Ampe, Gert Goeminne	**8F**	Interdisciplinarity in action: reflections from engaged research with chemical engineers	(this abstract should be considered in conjunction with the abstract by Goeminne and Ampe entitled: 'Unveiling narratives of innovation: A psychoanalytic take on engaged STS research')Interdisciplinarity in action: reflections from engaged research with chemical engineersKasper Ampe and Gert Goeminne There are increasing calls for socio-technical integration and inter- and transdisciplinary collaboration (Fisher, 2019). However, it is not clear how these expectations fit with the messy realities of technology-oriented R&D practices (Turnhout et al., 2020; Viseu, 2018). Guided by a framework called Socio-Technical Integration Research (STIR) (Fisher et al., 2006), this study aims to explore these messy realities through engagement with two chemical engineers. For twelve consecutive weeks, weekly dialogues were held with a full professor running a world-leading chemical engineering laboratory and a PhD student working in the same lab. The discussions, and of course their research, focus on the recycling of plastic waste into pyrolysis oil, an alternative to naphtha, and the electrification of the cracking process of this pyrolysis oil. This topic is also at the heart of a project involving the STIR researcher who organised the dialogues and the two engineers. The dialogues explored the choices the engineers face on a daily basis and the considerations, alternatives and decisions involved. On the basis of the framework and the empirical material, this study has, for now, the broad goals to understand (1) what societal considerations the engineers identify in their work, (2) how they intend to address these considerations, (3) what capacity they have to address them and (4) how this capacity might be improved? The study's preliminary observations for each of these four questions are as follows. With regard to the first and second questions, the two engineers tend to adopt rather standard techno-economic notions and practices. In particular, they believe that their technologies are a step in the right, sustainable direction. Society should accept these technologies, and policymakers need to address market formation and regulatory bottlenecks. But these notions and practices are surprisingly confusing, fragmented and inconsistent, which becomes particularly clear in the context of the third question below. This inconsistency emerged particularly in relation to the engineers' capacity to address the societal considerations associated to their technologies. Interestingly, they suggest that their capacity to do otherwise – which remained a vague and ambiguous notion during the dialogues – is constrained by established paradigms, expectations, networks, laboratory equipment and broader infrastructures. This situation leads them to adhere to their techno-economic notions and practices, as well as to push aside any responsibility to use their capacity to act, which bears some resemblance to previous findings (Lee et al., 2019). Regarding the fourth question, when the previous observations are combined with a secondary objective of the STIR framework, namely to enhance the capacity of engineers to address societal considerations, socio-technical integration researchers face a difficult task. Specifically, this study followed STIR in engaging with the local contexts of engineers, in particular avoiding the introduction of alternative contexts and perspectives. For now, the results of this study's intervention are rather sobering, perhaps requiring a move towards articulation, interpretation or assessment (Fisher et al., 2024) during the STIR intervention. The latter is explored from a psychoanalytic perspective in a companion abstract by Goeminne & Ampe. References Fisher, E. (2019). Governing with ambivalence: The tentative origins of socio-technical integration. Research Policy, 48, 1138–1149. Fisher, E., Mahajan, R., & Mitcham, C. (2006). Midstream Modulation of Technology: Governance From Within. Bulletin of Science, Technology & Society, 26, 485–496. Fisher, E., Smolka, M., Owen, R., Pansera, M., Guston, D., Grunwald, A., Nelson, J., Raman, S., Neudert, P., Flipse, S., & Ribeiro, B. (2024). Responsible innovation scholarship: Normative, empirical, theoretical, and engaged. Journal of Responsible Innovation, 11, 2309060. Lee, E. A., Gans, N., Grohman, M., & Brown, M. (2019). Ethics as a rare bird: A challenge for situated studies of ethics in the engineering lab. Journal of Responsible Innovation, 6, 284–304. Turnhout, E., Metze, T., Wyborn, C., Klenk, N., & Louder, E. (2020). The politics of co-production: Participation, power, and transformation. Current Opinion in Environmental Sustainability, 42, 15–21.
Lahari Yaddanapudi, Erik Fisher, Julia Hahn	**8F**	Socio-Technical Integration with early-career researchers in a large scale German COVID-19 project	In times of health crises like the COVID-19 pandemic, robust connections between science, society, technology and policy are not just desirable but essential. To achieve this, we must embrace interdisciplinary exchange methods that extend beyond traditional natural science realms, encompassing integrated social science approaches. Importantly, such collaborations can help enhance the reflective practices of scientists and researchers who are contributing to scientific advancement and technological development in times of crises to better align science and technology with policy and society. Among these integration approaches lies the Socio-Technical Integration Research (STIR) method, that has been extensively applied to assess scientific capacities for governing emerging technologies. However, its application in urgent health crises with evolving scientific evidence, like the COVID-19 pandemic, is yet to be explored. We applied the STIR method within a large-scale 5-year German project involving five early-career researchers focused on detecting and deactivating the COVID-19 virus in aerosols. The methodology encompassed pre-study interviews, 12 weeks of weekly decision protocol dialogues, and post-study interviews with each participant. Our analysis revolves around three key themes: Social Impact, Responsible Innovation, and Public Trust. With respect to the three themes, we aimed to identify the ideas and gaps as seen by the participants, and the study's outcomes manifest in three dimensions. Initially, drawing on the pre-study interviews, we identify values, considerations, and actions that participating scientists think are important. Subsequently, through the lens of the midstream modulation framework on the STIR dialogues, we unveil current practices. Finally, we assess potential advances and gaps by comparing what participants say during interviews and the concrete decisions they take after the interdisciplinary dialogues. Our assessments and recommendations are informed by reflexive discourses during the post-study interviews, supplemented by suggestions regarding areas necessitating further attention and prospective action. We found that the STIR dialogues documented, and in some cases brought about, changes in participating scientists' awareness of social context, awareness of social values they invoked, and attitudes and behaviour towards societal considerations. In particular, participants who did not initially see a direct social impact of their work pertaining to the COVID-19 pandemic became aware over time of important societal dimensions, as well as intended and unintended consequences of their decisions. For example, a participant who believed that "…there was no social impact on..." his work broadened his opinion to include the cost-effectiveness and health benefits of his technology on people, and expanded his vision to acknowledge conditions for its societal acceptance. This study therefore makes a threefold contribution. First, it underscores the importance of engaging scientists in discussions about the social impact of their work during crises with emerging scientific evidence, as a means to foster public trust. Second, it provides evidence that such integration enhances reflexivity among scientists and researchers. This suggests an augmented scientific capacity to voluntarily consider societal factors, thereby exhibiting a capacity for responsible innovation. Incorporating these discussions ensures responsible application of emerging knowledge for societal benefit, during crises like COVID-19. Lastly, it attempts to provide a broad framework to understand, assess and present the results of STIR studies.

Gert Goeminne, Kasper Ampe	**8F**	Unveiling narratives of innovation: A psychoanalytic take on engaged STS research	Unveiling narratives of innovation: A psychoanalytic take on engaged STS research "I consider you as my psychologist," a chemical engineer remarks jokingly at the conclusion of several weekly Socio-Technical Integration Research (STIR) dialogues. This paper takes this statement seriously and sparks a dialogue between Responsible Innovation (RI) and Psychoanalysis, exploring their shared focus on the role of narrative in processes of change. Grounded in co-productionist scholarship, RI advocates for considering the social dimension of research and innovation from the outset, in particular through engaged STS research. (Fisher 2006; Gjefsen et al. 2014). At this very point, where reflexivity regarding one's own position is associated with the unlocking of new possibilities, this paper proposes exploring parallels with psychoanalytic theory and praxis. Here, analyst and analysand, through structured interaction, seek to map an individual's narrative anchorage in their environment and render it actionable. Illustrated by a recent STIR study with chemical engineers (cfr. abstract by Ampe & Goeminne), this contribution explores the convergence of RI and psychoanalysis along two axes. Firstly, it parallels how both approaches center on narrative and self-discovery as catalysts for change. It therefore links recent interest within STS on the ontological role of narrative in responsible innovation (e.g. Reijers 2020, Sigl et al. 2020) to the linguistic turn in Lacanian psychoanalysis (Fink 1997). Consequently, it examines how R&D actors mobilize -often contradicting- individual, organizational and cultural narratives to inform and justify their daily research decisions. It thereby suggests understanding STIR's analytical framework of reflexive and deliberate modulations in terms of reconfiguring R&D actors' narrative repertoire. Secondly, it scrutinizes the working relationship between STS scholars and R&D actors in engaged STS research, likening it to a psychoanalytic 'talking cure.' It explores how psychoanalytic discourse theory and concepts such as (counter-)transference (Dulsster 2021) shed light on the diverse stances an STS researcher may adopt in establishing a productive relationship of 'generative critique' (Smolka 2020). In this respect, an intriguing distinction is found in the fact that a talking cure takes it starting point in the analysand's distress, a dynamic seemingly absent in engaged STS research. References: Dulsster et al. (2021) Lacanian discourse theory and the process of change in Lacanian-oriented talking therapies. Psychoanalytic Psychology 38(4): 319–327 Fink, B. (1997) Aclinical introduction to Lacanian psychoanalysis:Theory and technique. Harvard University Press Fisher, E. et al. 2006. Midstream modulation of technology: Governance from within. Bulletin of Science, Technology & Society, 26(6), 485–496 Gjefsen, M.D. et al. 2014. From Ethnography to Engagement: The Lab as a Site of Intervention. Science as Culture 23(3): 419-431 Reijers, W. (2020) Responsible innovation between virtue and governance: revisiting Arendt's notion of work as action. Journal of Responsible Innovation 7(3): 471-489 Sigl, L. et al. (2020) "I am primarily paid for publishing…": Narrative framing of social responsibilities in academic life sciences. Science & Engineering Ethics 26: 1569-1593 Smolka, M. (2020) Generative critique in interdisciplinary collaborations: from critique in and of the neurosciences to socio-technical integration research as a practice of critique in R(R)I. Nanoethics 14:1-19
Ordel Brown, Katherine Brichacek, Laura Pigozzi	**9A**	Operationalizing Empathy in Engineering Design via an Epistemic Tool to Support Equitable Design Outcomes	The demands of the ideological nature (e.g. social, cultural, environmental, political) of engineering design call for engineers to consider the broad societal implications of their work to include design equity and justice. From this perspective, the focus of engineering design moves beyond merely solving problems. There is the added focus of diagnosing and framing problems along dimensions of the human experience with an emphasis on the fundamental causes of the problems, and the pluralistic environments within which the problems arise. This emphasis undergirds humanistic considerations, such as empathy, within the technical domains of engineering design, and encapsulates empathic engineering as a critical precursor to design equity and justice. The main purpose of this paper is to describe an approach to empathic engineering that centers a multilevel and interdisciplinary pedagogical technique of operationalizing empathy in engineering design for equity and justice. The first level reviews the application of empathy in engineering design education through the humanistic-leaning lens of human-centered design (HCD), outlining benefits, shortfalls, and opportunities. The second level uses preliminary results from ongoing studies to address empathy-related shortfalls. Theoretically, this level combines philosophical empathy frameworks grounded in virtue and care ethics with HCD pedagogy to enable epistemic justice and facilitate equitable design outcomes. Practically, this level highlights a new direction for incorporating equity-focused empathy instruction in engineering design education. It describes pedagogical interventions that augment the largely cognitive and affective empathy emphases of HCD with conative empathy. This holistic empathy model allows for critical evaluations of mindsets, motivation, and methods throughout the design process and helps to maintain the design praxis within equity and pro-justice domains. The third level situates the interventions within the broader discussions of a reorientation of engineering design education and defines future work in creating more responsible futures. References 1. B. -A. Schuelke-Leech, "The Place of Wicked Problems in Engineering Problem Solving: A Proposed Taxonomy," 2020 IEEE International Symposium on Technology and Society (ISTAS), Tempe, AZ, 2020. 2. S. Costanza-Chock, Design Justice: Community-Led Practices to Build the Worlds We Need. MIT Press, 2020. 3. D. Bairaktarova, "Caring for the future: Empathy in engineering education to empower learning," Journal for Engineering Education, vol. 111, pp. 502-507, 2022. 4. C. Calloway-Thomas, "A call for a pedagogy of empathy," Communication Education, vol. 67, no. 4, pp. 495-499, 2018. 5. S. Afroogh, A. Esmalian, J. Donaldson, and A. Mostafavi, "Empathic Design in Engineering Education and Practice: An Approach for Achieving Inclusive and Effective Community Resilience," Sustainability, vol. 13, no. 7, p. 4060, 2021.
Anna Melnyk	**9A**	DESIGN FOR VALUE CHANGE AS A CLIMATE ACTION APPROACH	Value change is an ethical problem with practical implications in engineering and technological design. Many technologies and socio-technical systems quickly lose their relevance and contribute to, among others, the growing amount of e-waste which pollutes landfills. Whereas engineers face these issues more often these days, the ethical approaches to a technological design largely remain silent about the relation between the problem of value change and much-needed climate action through the design process. Some studies scrutinise the problem of value change in technological design, facilitate a long-term vision in multi-life span systems, and emphasizes the need to address extended time frames more explicitly. Yet, this scholarship doesn't systematically reflect the environmental concerns underlying the problem. In this presentation, I suggest taking a step further in claiming that the problem of value change should hold a more prominent position in the ethics of design discourse because it is inherently a climate action problem.

Nick Corvino	**9B**	Randomness as a Solution to the Self-Driving Car Dilemma	Abstract: Autonomous vehicles will inevitably encounter trolley car dilemmas, which may inhibit their path toward deployment even when they are safer than human drivers. If deployment is delayed by even a few years, this could result in the counterfactual deaths of thousands to hundreds of thousands of people (Kalra & Groves, 2017). Philosophers have proposed various solutions, often espousing a particular ethical theory such as deontology (Powers, 2006) or consequentialism (Anderson et al., 2004) in their defense. The problem, however, is that there has long been considerable disagreement about which ethical theory we ought to adhere to in trolley car scenarios. Given this moral disagreement, favoring one ethical theory and dismissing the others might be imprudent. Perhaps a more tractable strategy is to seek solutions that are morally inclusive. This paper offers a solution by proposing a 'stochastic variable' within autonomous vehicles. This feature gives representation to multiple ethical decision procedures based on a statistical distribution and selects one randomly when faced with a trolley car dilemma. Sometimes it will make the utilitarian decision, and at other times it will make a non-utilitarian decision. Sometimes it will sacrifice the safety of its passenger, and sometimes it will sacrifice the safety of an innocent bystander. Doing so aims to establish a decision-making procedure that gives representation to a wide range of ethical frameworks rather than favoring one and disregarding the rest. The stochastic variable approach highlights the need for an interdisciplinary effort, bridging philosophy and engineering. Non-philosophical questions demand our immediate attention, such as how to program randomness into the vehicle's decision procedure, how to assign precise weights for ethical representation, and considerations as to the cost and efficacy of this approach compared to viable alternatives. A multifaceted approach integrating philosophical rigor and technical expertise is vital to ensure autonomous vehicles uphold moral principles while operating safely on public roads. Neither field can solve this alone. References: Powers, T. (2006). "Prospects for a Kantian Machine." IEEE Intelligent Systems, 21, 46–51. Kalra, N. & Groves, D. (2017). "The Enemy of Good: Estimating the Cost of Waiting for Nearly Perfect Automated Vehicles." RAND Corporation. Anderson, M., Anderson, S. L., & Armen, C. (2004). "Towards machine ethics." AAAI-04 Workshop on Agent Orientations: theory and practice.
Sebastian Krügel, Matthias Uhl	**9B**	The global perspective on the risk ethics of autonomous vehicles	There is no road traffic without risks and any participation in road traffic inevitably entails a distribution of risks between road users. This raises ethical questions, especially when this distribution is enforced by autonomous vehicles (AVs). Patents from Google, Waymo and others show that the car industry is thinking about the risk management of AVs and that AVs allow for a more deliberate centralized allocation of risks between road users than impulse-driven decentralized manual traffic allows. The questions are which variables should be included in this risk management and how AVs should then select the supposedly appropriate maneuver from several possible ones. Building on a visual decision paradigm by Krügel and Uhl (2024), we conducted an international study in eight countries with a total of almost 11,000 participants to measure preferences of risk distributions in road traffic in different cultures. Based on the Inglehart-Welzel Cultural Map of the World from 2023, we selected a country from each of the eight country clusters for our survey (included countries: China, Germany, Greece, India, Mexico, Spain, South Africa, USA). Each participant in the survey was presented with one of seven possible traffic situations in which traffic risks could be distributed differently by choosing the safety distance of an AV between other road users. Reductions in traffic risks for some road users were accompanied by increases in traffic risks for others. In some scenarios, the participants themselves were part of the traffic situation, in others they were not. The most important result of our study is that the "social dilemma of autonomous vehicles" (Bonnefon, Shariff and Rahwan, 2016) in risky traffic situations does not occur in any of the eight countries surveyed. Contrary to deterministic dilemma situations, participants in our study do not expect AVs to minimize the risks for them as passengers at the expense of others. The second most important result is that cultural differences in the distribution of risk in road traffic are surprisingly small. The chosen safety distances and the associated risk distributions are very similar in all countries, in most cases statistically indistinguishable. Taken together, our study shows that population surveys on the driving behavior of AVs can be meaningful, even if the respondents themselves are part of the traffic situation, and that an international consensus on risk distributions in road traffic may indeed be possible. References Krügel, S., & Uhl, M. (2024). The risk ethics of autonomous vehicles: an empirical approach. Scientific reports, 14(1), 960. Bonnefon, J.-F., Shariff, A., & Rahwan, I. (2016). The social dilemma of autonomous vehicles, Science, 352(6293), 1573-1576.
Alexander Bagattini, Michael W. Schmidt	**9C**	Moral Expertise for Emerging Technologies	Moral expertise can be understood as the combination of a) knowledge of a wide range of ethical considerations, that can be gained, for example, by the study of ethical theories and b) a set of abilities and motivational sets that help identify moral issues and give adequate guidance in moral reflection and action. One might additionally demand that the corresponding disposition of such ethical experts is reliable in tracking ethical truth or correctness. Even if we assume that such ethical expertise is achievable, it is a further question whether there can be ethical expertise on emerging technologies. Ethical expertise on moral issues of emerging technologies would involve additional complications. Emerging technologies come with considerable uncertainties regarding their further development and implementation in society and consequently with uncertainty regarding their consequences. If there is ethical expertise on emerging technologies, then there are ways to deal with these uncertainties adequately. In this paper we argue for specific competencies/abilities and motivational sets that are at least necessary of moral expertise in emerging technologies.

Jon Rueda	**9C**	Techno-moral progress: Exploring the technological mediation of better morality	Moral progress and technological progress do not necessarily go hand in hand. The twentieth century is a prime example in this regard. According to various commentators (Mitcham 1994; Ihde 1990; Verbeek 2011), the atrocious two world wars and the growing environmental impact of technological societies spread among many postwar philosophers a critical view of modern technology. Certainly, material progress (mainly produced thanks to economic and scientific-technological advancement) does not equate to progress towards a more humane world. That said, the co-evolution between technology and morality is arousing increasing philosophical interest. The literature on "techno-moral change", in particular, is studying not only how morality influences technology, but how technology affects morality (Danaher & Sætra 2023; Hospter et al. 2022; Nickel et al. 2022; Swierstra et al. 2009). This co-shape between technology and morality may be interesting to attend to with regard to moral progress. However, research on techno-moral change has rarely analyzed this phenomenon in relation to the philosophical problem of moral progress, with some exception (Author 2024). In this talk, I offer a pioneering contribution about the technological mediation of moral progress. I thus intend to connect two growing philosophical literatures, that of techno-moral change and that of moral progress, which can benefit from studying their intersections. My main argument is that we should take seriously the possibility of techno-moral progress, given that human morality may genuinely advance through some technologically mediated changes. The structure of my talk is the following. First, I begin by motivating the topic and offering various conceptual clarifications. Then, I argue how we should understand the phenomenon of techno-moral progress, on the one hand, by avoiding teleological conceptions of moral progress, and, on the other hand, by suggesting the most plausible causal models of techno-moral changes. I then briefly discuss the issue of techno-moral regress, the other side of the coin of techno-moral progress. Finally, I address the limits of our techno-moral imagination in relation to future morality, in order to address some philosophical paradoxes that arise with respect to the alignment of moral values with future generations. References: Danaher, J., & Sætra, H. S. (2023). Mechanisms of techno-moral change: A taxonomy and overview. Ethical theory and moral practice, 26(5), 763-784. Hopster, J., Arora, C., Blunden, C., Eriksen, C., Frank, L. E., Hermann, J. S., Klenk, M. B. O. T., O'Neill, E. R. H., & Steinert, S. (2022). Pistols, pills, pork and ploughs: the structure of technomoral revolutions. Inquiry (United Kingdom), 1–33. https://doi.org/10.1080/0020174X.2022.2090434 Ihde, D. (1990). Technology and the lifeworld: From garden to earth. Indiana University Press. Mitcham, C. (1994). Thinking through Technology: The Path between Engineering and Philosophy. Chicago University Press. Nickel, P. J., Kudina, O., & van de Poel, I. (2022). Moral Uncertainty in Technomoral Change: Bridging the Expl anatory Gap. Perspectives on Science, 30(2), 260–283. https://doi.org/10.1162/posc_a_00414 Swierstra, T., Stemerding, D., & Boenink, M. (2009). Exploring Techno-Moral Change: The Case of the ObesityPill. In P. Sollie & M. Düwell (Eds.), Evaluating New Technologies: Methodological Problems for the Ethical Assessment of Technology Developments (pp. 119–138). Springer. Verbeek, P.-P. (2011). Moralizing Technology: Understanding and Designing the Morality of Things. Chicago University Press.
Ivano Zanzarella	**9D**	Mori's Uncanny Valley Historicized. The Epistemological Validity of HRI Quantitative Models on the Test-Bench of the History of Science	The theoretical reflection about engineering and technology has until now scarcely taken advantage of an approach integrating history and philosophy in the sense of, for example, &HPS (Integrated History and Philosophy of Science). Here, I introduce an Integrated History and Philosophy of Engineering and Technology (HPET), with particular reference to the problem of the epistemological validity of quantitative models of Human-Robot Interaction (HRI). Among the most known models of this kind, there is Masahiro Mori's [1970] Uncanny Valley. This model establishes a non-linear relation between human-likeness of artifacts' appearance and behavior, and emotional response of interacting humans. It moreover predicts that when the former is particularly high, the positivity of the latter drastically sinks –from which, namely, the name of the model itself. Although the phenomenon (arousal of uncanny feelings by highly human-like artifacts) is attested by large anecdotal testimony, and has been as such an object of scientific inquiry even since early 19th century (see e.g. Jentsch's [1906] studies on the Unheimlich), Mori's model is still considered a hypothesis. Most criticisms about its validity come from the empirical sciences (psychology, neurosciences, etc.). For instance, experimental evidence for it is in many cases just minimally replicable [Palomäki et al., 2018], and shows clear characters of inconsistency –e.g. both non-linear and linear versions of the model can be confirmed [Wang et al., 2015]. The alleged generality of the model, which should logically derive from its quantitative nature, does hence not hold, and its epistemological validity is undermined by its significant dependency on the subjects and the contexts considered within experimentations. Yet, not only is Mori's model epistemologically problematic from an empirical perspective, but also from a scientific-historical one: Its generality can be questioned, in fact, even with regard to a temporal-historical dimension. Indeed, by analyzing primary historical sources about the reception of 18th-century musical automata (in particular Vaucanson's Le flûteur, Jaquet-Droz' La musicienne and Kinzing/Roentgen's Die Hackbrettsspielerin), it can be noticed that, despite the predictions of the model, high degrees of human resemblance of such automata would only hardly (to never) elicit uncanny feelings in the interacting public of the time. In the true spirit of HPET, thus, epistemological and philosophical conclusions are drawn from these historical investigations: There cannot be any absolute "mathematical" relation between artifacts' human-likeness and human emotive response as alleged by Mori's model, for this relation is determined (also) by cultural and historical factors. These conclusions are then extended from Mori's model to every quantitative approach within HRI research: In general, deterministic and quantitative HRI models should be regarded as having a mere heuristic value. On the contrary, a study perspective on HRI is needed which is more aware of the social, cultural and historical factors playing a function in it beyond just the naturalistic, biological and mechanistical ones. References Jentsch, E. A. (1906). Zur Psychologie des Unheimlichen. Psychiatrisch-neurologische Wochenschrift, 22(23):195–198, 203–205. Mori, M. (1970). [Bukimi No Tani]. Energy, 7(4):33–35. En. Tr. MacDorman, Karl F. and Kageki. N. (2012), The Uncanny Valley. IEEE Robotics & Automation Magazine, 19(2):98:100. Palomäki, J. et al. (2018). Evaluating The Replicability Of The Uncanny Valley Effect. Heliyon, 4(11):e00939. Wang, S. et al. (2015). The Uncanny Valley: Existence And Explanations. Review of General Psychology, 19(4):393–407.

Colum Finnegan	**9D**	Engineering Online Communication Using Mindshaping Theory	Online communication technology is integral to the functioning of contemporary society. However the manner in which these digital affordances enable, structure and constrain communication possibilities is determined by design choices made by their architects. These choices are often guided by economic rationales, in particular the development of a captive and engaged user base. Understanding in specific how these choices may inadvertently generate pernicious outcomes for society is of crucial importance in order to better guide the engineering these technologies. In this talk I will show how recent research on social cognition in cognitive science can be put to use in understanding how exactly these technologies may be interacting with human communication in problematic ways. These new understandings of human social cognition are broadly termed interactionist accounts. Approaches of this nature foreground the role of interpersonal interaction in the development and practice of mindedness. The most empirically rigorous version of such an approach has come to be know as mindshaping. This is the idea that humans pervasively regulate one another to adopt shared and aligned cognitive profiles in order to enable interpersonal action coordination. Accounts such as this were developed explicitly in order to challenge the still dominant understanding of human social cognition known as mindreading. This is an explanatory paradigm that understands humans agents as atomistic truth seekers aiming to discover the mind states of others to enable action coordination. The innovation central to the mindshaping paradigm is to understand human agents as instead knitted into influence networks, pervasively regulating one another to adopt aligned cognitive profiles. Thus, according to mindshaping theory humans do not read one another in social contexts, rather they shape one another to have aligned, and thus relatively transparent mind states. This novel view of human social cognition has far reaching implications for our understanding of human communication. And as I will demonstrate, if it is on the right track, mindshaping has important implications in the context of optimal design of online communication affordances. Thus far a mindreading based account has implicitly informed the growing body of research into the pernicious effects of social media and digital communication affordances, leading to ineffective or counterproductive solutions being proposed by researchers. A perspective that instead understands humans as shaping one another to adopt certain belief sets can, I will argue, be used to better design online communication affordances in order to avoid pernicious outcomes. For example, the rise of intergroup polarisation online can be usefully explained from the mindshaping perspective, illuminating features of this phenomenon that mindreading based accounts miss.
David Goldberg	**9E**	Co-Contraries and Change in Higher Education	Since leaving the university to help drive change in engineering education (Goldberg & Somerville, 2014) and higher education more generally (Goldberg, 2023), some of the most useful hacks may be borrowed from the practice of executive coaching (Jirouskova & Goldberg, 2023). Among these, one of the most useful tools is the practice of framing change as the management of co-contraries. The paper starts by discussing what co-contraries are and connecting them to more familiar philosophical concepts. Co-contraries are related to Aristotle's idea of the golden mean as discussed in the Nicomachean ethic. For example, courage may be thought of as the balance of cowardice and foolhardiness, but as useful as this way of thinking is, the middling approach isn't the only way to manage more generalized co-contraries. Johnson (1992), using the term polarities, defines them as "opposites that need each other," but not all co-contraries are opposites and the phrase "need each other" has many possible interpretations. His example of breathing as the properly managed combination of inhaling && exhaling (we will use the symbol && to separate the terms of a co-contrary) suggests a kind of cyclical or dynamic management of co-contraries, which is often the case. The literature of paradoxes might be useful to us, but Reschler (2001), dismisses rhetorical juxtaposition of opposites as not a proper subject of philosophical inquiry. Moreover, he dismisses the rich and perfectly practical co-contrary of a general's responsibility to both protect && expose his troops in combat as mere irony. A more germane point of departure for the talk is to study Elbow's famous text Writing with Power (1981) and his somewhat less well-known text Embracing Contraries (1982). In the case of writing, Elbow's approach of creating freely && editing tightly is a practical way of managing these two different mindsets in a way that reliably generates high quality writing. Likewise, his addressing the fundamental contrary in teaching between a professor's authority and a student's freedom is in many ways the primary lesson of A Whole New Engineer and its recommendation of a system of education that unleashes lifelong curious learners by restraining authoritarian and dogmatic teaching and permitting substantially more student freedom and choice. From here, the talk examines many of the key co-contraries that are rife in educational change: individual && teamwork, authority && freedom, theory && practice, head && heart, self && other, planning && experimentation, and many more. The use of a sunshine-shadow quad chart to map out the benefits of adequate emphasis of a contrary and difficulties that arise with over-emphasis is also demonstrated. The talk closes by discussing some of the ways in which co-contraries can be managed through the intentional modification of personal habits, group rituals, values, and organizational structures and culture. References Elbow, P. (1981). Writing with power: Techniques for mastering the writing process. Oxford University Press. Elbow, P. (1982) Embracing contraries: Explorations in learning and teaching. NOxford University Press. Goldberg, D. E. (2023). A field manual for a whole new education: Rebooting higher education for human connection and insight in a digital world. ThreeJoy Associates, Inc. Goldberg, D. E. & Somerville, M. (2014). A whole new engineer: The coming revolution in engineering education. ThreeJoy Associates, Inc. Jirouskova, J. & Goldberg, D. E. (2023). How modern coaching can help develop engineers and the profession—and how philosophy can help. In A. Fritzsche & A. Santa-María (eds.), Rethinking technology and engineering: Dialogues across disciplines and geographies. Springer Verlag. pp. 81-99 Johnson, B. (1992). Polarity management: Identifying and managing unsolvable problems. Amherst, MA: HRD Press. Rescher, N. (2001). Paradoxes: Their roots, range, and resolution. Chicago, IL: Open Court.

Andrés Santa-María	**9E**	Contemporary Philosophy of technology and its pedagogical opportunities	As Mitcham (2019) said, it is not the same to talk about a philosophy of engineering and philosophy for engineers. The former refers to the wide range of topics that has to do with the reflection on the philosophical fundamentals and implications of engineering knowledge and practice. The latter, instead, refers to what engineers can get from philosophy (for example during their studies) to be better engineers. Indeed, not every topic addressed as part of the general discussion held as philosophy of engineering has the same formative potential for a prospect engineer. This distinction is particularly relevant if it is considered in the general context of engineering education, where there is a tension among institutional discourses that progressively recognize the importance of a sustainable and reflective engineering practice and a curriculum that still resist to incorporate core subjects beyond the scientific and technical disciplines traditionally considered as the essence of engineering knowledge. Such situation means that, in most of the cases, engineering students used to having just a few instances in their programs (when they have any) to focus on developing tools to think critically about engineering, in a way that may help them to be better engineers. In this paper I will address the problem of how to define what a philosophy for engineers may be, especially when some of our resources (time, credits) are scarce. In other words: how to set the "philosophical must-learns" for a prospect engineer? As a possible answer to this question, I will suggest that the philosophy of technology developed since its so-called "empirical turn", about 30 years ago, is an important fount of fruitful philosophical discussions for engineers. I will focus particularly on the pedagogical opportunities of the "theory of technological mediation", developed, among others, by Verbeek (2011). This theory seems to be, in the first place, a well-balanced approach to engineering ethics (between the pessimism of a preventive approach and the optimism of an aspirational one; Kanemitsu 2018). In the second place, it may help a deeper understanding of the relation between ethics and engineering, i.e. not as a mere frame that engineering must not surpass, but as the starting point of responsible innovation, as has been noted by the latest works on the idea of "design for values" (van de Poel 2015). Finally, the understanding of the mediating role of technology in human life may support a better consideration of the many variables that constitute the kind of problems that engineering used to dealing with, including the ethical/political variables in their very weight. That means that this approach can contribute to a better understanding of the concept of "engineering problem" and, thus, of engineering itself. References Kanemitsu, H. (2018). New Trends in Engineering Ethics–A Japanese Perspective. In The Future of Engineering (pp. 243-256). Springer, Cham. Mitcham, C. (2019). Steps toward a philosophy of engineering: historico-philosophical and critical essays. Rowman & Littlefield Publishers. van de Poel, I. (2015). Design for values in engineering. Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains, 667-690. Verbeek, P.P. (2011) Moralizing Technology: Understanding and Designing the Morality of Things. University of Chicago Press
Sabine Ammon	**9F**	Ethical Co-Design for Responsible Technology Futures: On the epistemology and ethics of integration in research and development	This contribution explores how integrated ethics can close the principles to practices gap in product development, which is currently seen as a major obstacle for ethics integration in innovation processes. By drawing on an example of an ethics intervention featuring an AI technology, I will show how ethics can become a co-designer for future products. I will argue that a successful integration process in research or development needs to embed ethical considerations into engineering reasoning and thinking, which, in turn, asks for an epistemology (and ethics) of integration. In processes of ethics integration, different ways of thinking and reasoning about technology collide. Integrated technology ethics uses this contrast to make ethical problems in technical design visible. Recognizing that technology and artefacts can be thought of in different ways is already an important first step towards examining ethical aspects in a development process. In a loose reference to Ludwig Fleck, I would like to address the collision of different ways of thinking about technology as different thought styles. The engineering perspective is characterized by a solution-oriented approach. It aims at a concrete result within a limited time frame, which is good enough; thus, efficiency and the reduction of complexity become epistemically relevant. In contrast, ethics is about answering more general questions of how we can live well together as human beings. Technology ethics expands this question to include the good coexistence of people under the impact of technology. The question is not about the interests or needs of an individual person, but the answers are aimed at a supra-individual validity. This means that thinking and reasoning in ethical analysis move from the concrete to the abstract in order to develop answers that can be applied to specific individual cases. I will argue that it is not enough to add a few values to the development process to make them ethically robust. The thought processes of both ways of grasping and processing technical objects are very different, and, at the same time complement each other. In integrated ethics, an interdisciplinary generation of knowledge takes place in a joint reflection space while assessing possible consequences and developing ethically desirable development alternatives. Apparently, this shift in reasoning and thinking is not easy to achieve and requires a special protective space, which leads us to questions of the ethics of integration. Acceptance of the divergent epistemic perspectives, trust in expertise, the security of being able to safely address potential problems, the ability to point out difficulties and mistakes, appreciation and dialog at eye level, all indicate how important process design and the culture of exchange are for integrated ethics. Once the integration framework is established, it becomes possible to ethically reframe the design problem and make it accessible for an ethical assessment. By comparing integrated ethics to other approaches for early intervention like embedded ethics, STIR, CTA/ real-time TA, disclosive ethics, VSD, vision assessment, I will conclude by discussing strengths and limitations of the integrated ethics approach.

Tilke Devriese	**9F**	Productive misunderstanding in multidisciplinary design: the influence of different epistemic backgrounds on model comprehension	Considering design as an epistemic practice (Ammon 2013) and design models as epistemic tools (Eckert and Hillerbrand 2018; Boon and Knuuttila 2009), necessitates to scrutinize the influence of varying epistemic backgrounds of collaborating designers in a multidisciplinary design process. This paper explores 'productive misunderstanding' between architects and structural engineers in design processes of several contemporary buildings, by looking at the use and comprehension of several design models during a collaborative design process. The examined design models are sourced from symmetrical reconstructions of several lengthy, messy design processes of contemporary architectural projects. These reconstructions are based on collaborating designers' design documents. Design documents - sketches, plans, notes,... - and the models within are fundamental tools through which design is enabled. On the one hand, this is where ideas are externalized and reificated from 'the mind's eye', which allows them to be 'verified'. On the other hand, design documents and models are the primary means to communicate design ideas (Henderson 1998). Models enable associative reinterpretation and productive misunderstanding, which makes them crucial media of a creative dialogue (Oswalt 2015). When considered in the context of multidisciplinary design, design documents can become boundary objects (Carlile 2002) - objects that allow collaborators to read different meanings particular to their needs from the same material. However, design models themselves are tools shaped by the epistemic culture in which they are produced (Knorr-Cetina 1999). Often specific knowledge is needed to fully grasp the content of a design model, which is why their comprehension by collaborators in multidisciplinary design does not fully coincide. The paper explores several examples of how exactly this 'misfit' in model comprehension can be activated and argues that the nuanced misunderstanding of models due to varying epistemic backgrounds of collaborators, can be a main propeller in the negotiation of a multidisciplinary design process. references: Ammon, Sabine. 2013. "Wie Architektur entsteht. Entwerfen als epistemische Praxis." In Wissenschaft Entwerfen: Vom forschenden Entwerfen zur Entwurfsforschung der Architektur, edited by Sabine Ammon and Eva Maria Froschauer, 336-361. Leiden, Netherlands: Brill | Fink. Boon, Mieke, and Tanja Knuuttila. 2009. "Models as Epistemic Tools in Engineering Sciences." In Philosophy of Technology and Engineering Sciences, edited by Anthonie Meijers, In Handbook of the Philosophy of Science., 693-726. Amsterdam: North-Holland. Carlile, Paul R. 2002. "A pragmatic view of knowledge and boundaries: Boundary objects in new product development." Organization Science 13 (4): 442-455. Eckert, Claudia, and Rafaela Hillerbrand. 2018. "Models in Engineering Design: Generative and Epistemic Function of Product Models." In Advancements in the Philosophy of Design, edited by Pieter E. Vermaas and Stéphane Vial, 219-242. Cham: Springer International Publishing. Henderson, Kathryn. 1998. On Line and On Paper. Visual Representations, Visual Culture, and Computer Graphics in Design Engineering. Massachusetts: The MIT Press.
Yuying Sun	**11A**	Ethical Considerations in the Construction of Digital Deceased People	Digital deceased is a kind of digital person, referring to the digital body of the deceased. The construction of digital deceased is to upload the text, pictures, sound, video and other digital information left by the deceased, and use digital technology to simulate the words, behaviours and personalities of the deceased, so that the deceased can continue his/her existence in the virtual environment in the form of a digital person and a digital body. The existence of the digital deceased can help the living and the deceased to establish communication and interaction and rebuild their relationship narrative. The rationality of the construction of digital deceased can be examined from the perspective that the construction of digital deceased meets the emotional needs of the living and cares for and remembers the dead, and the existence of digital deceased can meet the needs of the living and the dead to keep in touch with each other, give spiritual comfort to the living, and comfort and heal the emotional trauma produced by the living due to the dead. The existence of the digital deceased provides a digital space for expression and remembrance, carrying the digital memory of the deceased, as if achieving a certain symbolic immortality, giving the living the opportunity to care for, commemorate and remember the deceased, so the digital deceased construct has a certain degree of rationality. In essence, what is constructed by the digital deceased construction is the digital identity of the person, and the identity provides a new perspective to examine the legitimacy of the digital deceased construction: the important symbol of the identity is the face of the person, and the digitally generated face of the deceased may cause the living to fall into the Valley of Terror effect; the digital deceased construction can only reproduce the original digital identity of the deceased, while the digital information that constructs the digital identity of the deceased may not be comprehensive and reliable. Moreover, the construction of digital deceased may go against the wishes of the deceased, and the use of the deceased's digital information and facial images may infringe on the deceased's privacy and portrait rights. The construction of digital deceased has its own rationality and needs, but the premise of constructing digital deceased is to respect the will of the deceased, and the construction can only be carried out with the expression of consent; individuals should do a good job in the management of the digital heritage; the protection of the deceased's privacy and other rights; and the guidance of the development of the technology of constructing the digital deceased to the good.
Ria Ariani	**11A**	Mapping Ethical Dynamics in Open Science: Data Sharing Challenges in Indonesian Qualitative Society	Open science has revolutionised the academic landscape, facilitating the dissemination of scientific findings and data to a wider audience. However, the implementation of open science practices often encounters ethical and legal complexities, particularly regarding confidentiality and data sensitivity. As technology advances, ethical considerations become even more crucial, especially concerning the sharing of qualitative data on open platforms.This study investigates the prevalent patterns and trends of ethical dilemmas arising from the sharing of qualitative data on open platforms, focusing on the Indonesian research context. Employing thematic analysis from ethical guidelines in Indonesia, the findings reveal the critical examinations of the adoption and the use of open science technologies. The main ethical issues in open science area concerning informed consent, privacy protection, and data ownership. Additionally, the complexities of cultural and societal contexts must be considered when implementing ethical guidelines for data sharing in diverse settings like Indonesia. Keywords: Open science, data sharing, qualitative data, ethical dynamics, Indonesia.

Yuqi Peng	**11A**	Unveiling AI-Induced Vulnerabilities: The Case of Deepfake Technology	This paper critically examines the significant risks and ethical quandaries that artificial intelligence (AI) technologies present, with a special focus on their impact on vulnerable populations and the generalized vulnerability phenomenon. It identifies and analyses vulnerable groups of AI to demonstrate their heightened vulnerability amidst AI's swift progression and advocates for tailored protective measures to shield them. The study suggests that AI's evolution may induce a widespread vulnerability phenomenon, thus establishing the universal pertinence of AI-related risks. Specifically, it delves into the ethical complexities and direct hazards posed by Deepfake technology, underlining the particularly severe risks for vulnerable communities. The paper introduces a blend of policy and technological interventions aimed at curbing the detrimental impacts of AI and prioritizes the protection of these at-risk groups. It underscores the essential need for ethical awareness among tech professionals and highlights the value of educational initiatives designed to fortify vulnerable populations against AI threats. By forwarding these arguments, the paper contributes to the ongoing dialogue on AI ethics, aiming to promote the fair distribution of AI's advantages while protecting the most susceptible segments of society.
Clint Hurshman	**11B**	Robust Social Goods: On the Value of Informational Privacy	This paper argues that informational privacy gets its distinctive value in virtue of making it possible to enjoy robust social goods. Information and communication technologies (ICTs) like social networks, facial recognition technology, and big data analytics have drastically restructured the social world in recent decades and continue to do so as more spheres of life become technologically mediated. While this trend has raised concerns from privacy theorists, there is still no consensus on exactly what makes privacy valuable. In particular, while philosophers agree that it protects a variety of goods such as safety, autonomy, and freedom, they disagree on whether its value is reducible to these goods, or whether it has value distinct from them. This paper elaborates on James Rachels' (1975) argument that privacy is necessary to maintain a variety of relationships, since cultivating varying degrees of intimacy requires revealing oneself differentially to others. Drawing from social ontology, I expand on this view by pointing out that not only close relationships, but a variety of social goods associated with them including respect, trust, and power consist (at least partly) in the mental states of those around an agent, and that these mental states often depend, in turn, on the (true or false) information about an agent that others possess or lack. For example, in order to occupy social roles such as that of a mentor, it's necessary that some "friction" (Floridi 2005) opposes the flow of "embarrassing" (i.e. role-undermining) information about an agent to her mentees. An unregulated flow of information makes social goods associated with being a mentor, such as respect and trust, unattainable or fragile. Privacy, broadly understood as implying a limitation on the flow or accessibility of information about a subject, thus makes such social goods possible and durable. The paper argues that this is the distinctive value of privacy because, while other resources can allow one to maintain autonomy, freedom, and safety even in the absence of privacy, the social goods under consideration are not possible without others occupying certain mental states, and thus not robust without limitations on the flow of information that could change those mental states. There is no substitute for privacy as a protector of robust social goods. The paper then considers some upshots of the argument for the ethics of technology. On one hand, the permanence and shareability of digital information, and the crossing of contexts that often occurs online (see Nissenbaum 2019), may straightforwardly undermine our ability to enjoy these goods by making information more accessible to those on whom our social status constitutively depends. On the other hand, by changing the nature of social relationships, the changing infosphere may prompt us to rethink the meaning of the goods under consideration. References Floridi, L. (2005). The ontological interpretation of informational privacy. Ethics and information technology, 7, 185-200. Nissenbaum, H. (2019). Contextual integrity up and down the data food chain. Theoretical Inquiries in Law, 20(1), 221-256. Rachels, J. (1975). Why privacy is important. Philosophy & Public Affairs, 323-333.
Katherine Brichacek	**11B**	Teaching effective altruism in an age of deepfakes	Effective altruism is the utilitarian "ethical" movement currently taking Silicon Valley by storm. I argue that effective altruism neglects the harms against women+ and girls by favoring oversight of future technology—not present technology. I take up the example of deepfakes to argue that effective altruism distracts us from the real—instead of possible future—harms against women+ and girls that advances in technology currently cause. Effective altruism ("EA") is a 21st century utilitarian attempt at maximizing good in the world. A few groups began to separately grow in popularity in the early 2000s including the Singularity Institute (now the Machine Intelligence Research Institute or MIRI), GiveWell (now Open Philanthropy), and the Centre for Effective Altruism (CEA) at Oxford. Additionally, philosophers Peter Singer and Will MacAskill have become persuasive faces of the EA movement which, while still operating via separate organizations has become significantly influential in Silicon Valley among tech moguls, engineers, and AI-based companies. MIRI and Open Philanthropy, in particular, have outsize influence over AI-based startups, often opting to only fund the startups that agree with their effective altruist theories. The Centre for Effective Altruism aims to "pay attention to how many people are effected by a particular invention" —a straightforward and worthwhile utilitarian aim in technology [1]. Yet, more alarming is their claim: "New technologies might threaten to wipe out life on earth, and reducing this risk might be a key priority" [1]. Also known as existential risk, the idea that advancements in technology—particularly AI—pose a serious and likely threat to the existence of humans is a hallmark of the effective altruism movement and one that I claim alienates those who are currently harmed by advances in AI technology. Deepfakes, or the artificially altered videos that falsely paste a different head or body onto an existing video, are a current risk to the online community of women+ and girls. 90-95% of deepfakes are of women+ and girls, pornographic in nature, and nonconsensual [2]. Open source deepfake technology such as DeepNude which "strips" the clothes off the picture of a person has affected roughly 104,852 women as of 2020 and saw a 138% increase in use over a three-month period in the same year [2]. The current serious harm of nonconsensual pornographic deepfakes effects women+ and girls and their ability to exist safely in online communities for fear that sharing clothed, mundane photos will be manipulated by deepfake technology. By exploring the gendered harms of deepfake technology, I argue that EA neglects the current effects of AI and, in turn, alienates women+ and girls in online communities. I take up this analysis of the harms enacted by EA to make a larger argument about how best to teach utilitarianism, alongside other ethical systems, in engineering ethics courses. Emphasis on the present implications of ethical decision making must be weighed in addition to future ethical considerations. [1] https://www.effectivealtruism.org/ [2] H. Adjer, G. Patrini, F. Cavalli, "Automating Image Abuse," Sensity, October 2020. Accessed: Feb. 12, 2024 [Online]. Available: https://www.medianama.com/wp-content/uploads/Sensity-AutomatingImageAbuse.pdf

Joost Mollen	**11C**	Experiments without Borders: Research Ethics, Real-World Experimentation, and the Identification Problem	Research participants cannot always be identified in real-world experiments with emerging technology. However, executing ethical obligations that render research morally permissible often requires an investigator to know who the research involves. These obligations include but are not limited to informed consent, the right to withdraw, weighing the research's benefits and risks to participants, compensating or remedying research-related harm, containing adverse consequences, debriefing participants, ensuring a just distribution in participant selection, and offering extra protections to vulnerable participants. The inability to fulfill these obligations presents a severe moral challenge for all researchers engaged in real-world experimentation. In this paper, I connect a fragmented debate on this issue and offer a descriptive account of the origin, scope, and consequences of this phenomenon, which I call the identification problem. I discuss the implications and significance of this problem to the growing field of real-world experimentation and outline ways forward. Without overwriting reasons, the identification problem renders real-world experiments morally impermissible. This prompts the need for further scholarly reflection on developing custom research ethics guidelines, adapting current real-world experiment practices, and developing research ethics governance mechanisms for public and private parties.
Zehua Yu	**11D**	Why does similarity works in Engineering?	There are many methods to understand the world, especially in science and engineering. Among these methods, experimental practice is paramount. Sometimes, to reduce costs, simulating the processes of scientific and engineering objects is crucial for understanding the world, which stimulates the creation of models, particularly scale models. Similarity plays a significant role in representing the world through scale models, relating the model and the world in specific aspects. It is so widely used that its processing logic and the doubts surrounding similarity need clarification. Only through rigorous questioning and effective justification can the concept of similarity be solidly grounded both philosophically and logically. We aim to refine the notion of similarity to uncover its potential value to engineers. The article first describes the evolution of the concept of "similarity," which has been distinct in historical trajectories and disciplinary variations in its interpretation and application. Then, it identifies certain attributes to describe similarity more logically and rigorously, comparing their advantages and disadvantages. Following this, we review several critiques of the concept of similarity, acknowledging their merit but noting they are insufficient to refute the concept entirely. We also highlight some issues that similarity cannot address, which many philosophers overlook. Based on this, we introduce the notion of "emphasized similarity", which means in different situations,Engineers will make different emphasize focus to make similarity more useful,after that we compared it with the previously identified attributes.
Emma Cavazzoni	**11D**	Sharing data, sharing technologies. Data-technology communities in Haly.Id	This talk explores the concept of data-technology communities as a unit of analysis to understand the emergence of knowledge-producing groups centered around engagement with common data collection technologies (Ankeny and Leonelli 2016). Data-generating tools can serve as meeting points for interdisciplinary collaborations, bringing people together in three interconnected ways. First, individuals converge around the technology itself: different people use a common instrument for data production. Second, group members share the same data generated by the common technology (Gregory et al. 2020; Springer and Cooper 2020). Third, participants experience the relationship of mutual influence that exists between data and technologies, whereby the tools adopted by researchers impact data collection, measurement, analysis, and outcomes (Kitchin 2014; boyd and Crawford 2012; Borgman 2015), and the type of data researchers aim to acquire shapes the technologies employed for their generation (Ribes and Jackson 2013; Poirier 2022). I thus define data-technology communities as heterogeneous knowledge producing groups of individuals whose essence consists in collectively dealing with shared technologies, data, and the reciprocal influence between them. This notion facilitates understanding how communities are formed as a result of interactions with specific technologies, and investigating the nature of scientific knowledge generated. The aim is to disclose community-specific standards of evidence and best practices around technologies, which shape the generation of scientific knowledge. In other words, by examining how individuals converge around common data collection technologies through the three interconnected means outlined above, I explore the establishment of standard of evidence and good practice that lay the groundwork for the scientific knowledge produced. I ground my reflections on a collaboration with the Haly.Id project in northern Italy, composed of a diverse group of people that gathers and deploys a high volume of data through remote sensing to limit the damage inflicted on crop fields by the brown marmorated stink bug Halyomorpha halys (H. halys) – a highly invasive pest that feeds on fruits and harms production (Bariselli, Bugiani, and Maistrello 2016; Ferrari et al. 2023). I show that Haly.Id harbors data-technology communities within it, built around some of the instruments adopted. The community centered around near-infrared hyperspectral imaging (NIR-HSI), for instance, includes some Haly.Id participants, bringing together individuals with varied expertise (e.g., analytical chemistry and engineering) who collaborate, share materials (pears), exchange insights, and co-publish findings. This highly heterogenous group of researchers engage with the same data collection technologies – NIR-HSI –, with the data generated through it – hyperspectral images of pears –, and with the relationship of mutual influence that exists between the two – NIR-HSI is modified and adapted to suit the acquisition of images displaying punctures on pears, and, in turn, it is tailored to capture specific data. According to my characterization, it thus qualifies as a data-technology community. Scientific knowledge developed in the NIR-HSI community regarding H. halys damage on pears relies on evidence that respect certain characteristics, such as images processed and cleaned to exhibit only fruits' features and excluding non-pear elements (e.g., petioles and labels), and on best practices around NIR-HSI, like rotating photographers at predetermined intervals. Bibliography Ankeny, Rachel A., and Sabina Leonelli. 2016. "Repertoires: A Post-Kuhnian Perspective on Scientific Change and Collaborative Research." Studies in History and Philosophy of Science Part A 60 (December): 18–28. https://doi.org/10.1016/j.shpsa.2016.08.003. Bariselli, Massimo, Riccardo Bugiani, and Lara Maistrello. 2016. "Distribution and Damage Caused by Halyomorpha Halys in Italy." EPPO Bulletin 46 (2): 332–34. https://doi.org/10.1111/epp.12289. Ferrari, Veronica, Rosalba Calvini, Bas Boom, Camilla Menozzi, Aravind Rangarajan, Lara Maistrello, Peter Offermans, and Alessandro Ulrici. 2023. "Evaluation of the Potential of Near Infrared Hyperspectral Imaging for Monitoring the Invasive Brown Marmorated Stink Bug." Chemometrics and Intelligent Laboratory Systems 234 (January): 104751.
Giovanni Frigo	**11E**	The Caring Engineer	This article explores the identity and role of engineers through the lens of care ethics. Building on other attempts to develop engineering ethics based on normative moral perspectives that are similar to and consistent with the ethics of care (e.g., virtue ethics), we propose that engineering practices within small and medium scale projects present the conditions for developing moral relations based on care. While in reality caring may partake in the activities and goals of the members and stakeholder within an engineering project, there is not much research about an engineering ethics that is explicitly grounded in care ethics. Moving from a characterization of care as "clusters" of practices and values (Held, 2006) (36, 40), the practices of caring actions resembles the exercise of virtues in that it requires availability, intention, and repetition. To illustrate the practical outcome of such a care engineering ethics, we provide a thought experiment regarding a small-scale energy project where the different caring virtues are exercised by the engineers involved.

Derek Schuurman	**11E**	Design Norms and Virtues for Engineering and Computer Science Education and Practice	This presentation will describe the use of "design norms" in the education of undergraduate engineering and computer science students. These design norms, derived from the theory of modal aspects developed by the Dutch philosopher Herman Dooyeweerd, include cultural, lingual, social, economic, aesthetic, justice, ethical, and trust aspects. The design norms will be compared with the Value-Sensitive Design framework with examples of how design norms can inform design in engineering and artificial intelligence (AI), These inclusion of design norms in the undergraduate engineering and computer science curriculum will be described, beginning in introductory courses and culminating in senior projects in which students must include a reflection on normative design considerations. In addition, it will be shown how design norms can enhance the teaching of engineering ethics. A traditional approach to teaching ethics relies on case studies which students address using traditional ethical frameworks such as deontological ethics, utilitarianism, or professional ethical codes. Such case studies often focus on specific dilemmas that often seem remote from everyday technical work. In contrast, design norms provide a broader and deeper approach that applies more to everyday engineering. Finally, it will be argued that teaching design norms is not sufficient on its own to train responsible engineers and computer scientists. Students in technical disciplines need to cultivate good character and virtues. Without developing virtue, the teaching of ethics has limited effectiveness. It is virtue that will incline an engineer to discern and respond appropriately to the norms that arise within a particular context. These virtues include many of the "technomoral" virtues suggested by Shannon Vallor, such as humility, courage, self-control, justice, empathy, care, civility, flexibility, and wisdom. Some concluding thoughts will be shared about practices that could be introduced within the engineering and computer science curriculum to help cultivate virtue in students. Selected References: Marc J. de Vries and Henk Jochemsen, eds., The Normative Nature of Social Practices and Ethics in Professional Environments, IGI Global, 2019. Ribeiro, P.F., Verkerk, M.J., Salles, R.S., "Toward a Holistic Normative Design" in Zambroni de Souza, A.C., Verkerk, M.J., Ribeiro, P.F. (eds) Interdisciplinary and Social Nature of Engineering Practices. Studies in Applied Philosophy, Epistemology and Rational Ethics, vol 61. Springer, 2022.
Pieter Vermaas	**11F**	Black Boxing Quantum Technologies in a Value-Sensitive-Design Exploration of Security Threats in the Port of Moerdijk	In this paper I present results of a Value Sensitive Design exploration of the security threats of future applications of quantum technology in ports and discuss the need that in such explorations stakeholders have basic knowledge of quantum technologies. The exploration was done with personnel from the Port of Moerdijk, which is a seaport in the Netherlands, through a workshop using the Security Threat Discovery Cards deck. The context of this paper is given by the challenges that come with the evaluation of the ethical and societal impact of quantum technologies. These challenges are that (1) applications are not yet well-developed thus obstructing detailed analyses, and that (2) quantum technologies are often perceived as incomprehensible. These challenges suggest that before evaluating the impact of quantum technologies, concrete applications should be found, and that stakeholders have to understand at least the basics of quantum technologies(e.g., Coenen and Grunwald 2017, Vermaas 2017). The result of the presented exploration suggest that these two preconditions need not be met for engaging into a productive analysis of the impact of quantum technologies. The Security Threat Discovery Cards enabled personnel of the Port of Moerdijk to analyse applications of quantum technologies and identify how it may be used, abused and compromised in a port. In the paper I present the application of quantum technologies given in the workshop, describe the security risks identified through the workshop, briefly compare these results with the risks identified by experts in port security (which acted as benchmarks to the workshop), and then expand on the need that stakeholders have a basic knowledge of quantum technologies in assessment of these emerging technologies. References Coenen C, Grunwald A (2017) Responsible research and innovation (RRI) in quantum technology. Ethics and Information Technology 19: 277–294 Security Threat Discovery Cards deck: https://securitycards.cs.washington.edu/ Vermaas P (2017) The societal impact of the emerging quantum technologies: A renewed urgency to make quantum theory understandable. Ethics and Information Technology 19:241–246

Alok Srivastava	**11F**	Tracing Responsiveness of Design Change to Value Changes in the Twelve Year History of a New Technology/Product - Semaglutide (Ozempic) Biopharmaceuitcals	In their new manuscript titled "Value Change Sensitive Design - Elements of a Process Ontological Framework and Method" which is being published by Springer in fPET 2023 Handbook of Engineering and Value Change, Srivastava and Bombaerts formulated a process ontological framework to enable methods for discovery of causal mechanisms underlying changes in values and other core elements of sociotechnical systems that affect design and outcomes of technology projects. Their approach addresses the need to monitor and respond to changes in values and other core elements during a technology project/product that are organized into the design of the project/product. They used Process Ontology and Niklas Luhmann's Social Systems Theory to develop a minimal ontological system and a diagramming tool – a schema - for tracing a minimal temporal change process centering on an 'event' in which a change is mobilized. This ontological framework and diagramming tool of contingent changes is combined with the tools of historical analysis called Path Dependence Analysis and Process Tracing to propose a full method for discovery and characterizations of causal mechanisms underlying change in sociotechnical systems and social processes. It is a contribution to the philosophical literature in Values Sensitive Design and Values Change. In this current paper I build on the work of Srivastava and Bombaerts by applying the framework and tools to the case of Semaglutide (Ozempic) - a drug developed initially for diabetes management and which has become popular for weight management. I will detail out the 12 year history from the formulation of the concept of the drug and its project of development through its approval by the FDA in 2017 and the following six years of its deployment and use in society. In this 12 year history I count at least four significant events that mark changes in each member of the constellation of the socio-technical system - i.e. values, stakeholders, choice-making and the technology/product. The four events are: 1. Conception and construction of the project to develop the technology/product in June 2012; 2. Approval of the drug by the FDA in October 2017; 3. Widespread recognition of its unprecedented impact in management of diabetes for a large number of patients by 2019; 4. The precipitation of the shortage of supply because of widespread prescription of the drug to non-diabetics for the purpose of weight-loss. In each of these events more than one of the four members of the constellation of the socio-technical system underwent change and these changes impacted the next event. I will use the Process Ontological framework and Process Tracing Methodologies to carry out causal analysis of the change process. My aim is to develop an illustration of the application of this approach and contribute to the methodological literature in Values Sensitive Design and Values Change in Technology and Engineering.
Anna Melnyk, Eugen Popa	**12A**	The Utopian Search for Energy (In)justice and The Need for Moral Pluralism	The concept of energy justice has emerged as a powerful tool for bringing political ideals into the complex study of energy transitions. With its "seeming universal appeal," the quest for justice and the counterpart quest of injustice(s) is an offer one cannot refuse. Methodologically, however, the approach suffers from both justificationism and utopianism, a combination that affects equally the descriptive reconstructions of current status and the normative prescriptions of future directions. It neither offers insights into navigating value conflicts nor reflects imminent value change. In this paper, we argue that the energy justice approach needs an infusion of moral pluralism to escape the utopian chase of perceived ideals (or perceived dangers), and advance its approach to dealing with value conflicts and potential value changes. Our argument is two-pronged. On a theoretical level we show that moral pluralism covers some of the problems previously identified with the energy justice approach but often only implicitly; to build our argument in favor of moral pluralism we identify potential ethical dilemmas value conflicts and value changes may pose. On a practical level, we illustrate how moral pluralism can be explicitly applied to scrutinize current developments in the hydrogen transition (the move from fossil-based to 'clean' hydrogen), a transition that has recently caught the attention of energy justice scholars, and what insights about value conflicts and value changes such application has to reveal. We conclude with a list of practical action points for consolidating the marriage between energy justice and moral pluralism.
Elisabeth Shrimpton, Nazmiye Balta-Ozkan	**12A**	OPERATIONALISING JUSTICE INTO INFRASTRUCTURE ENGINEERING RESEARCH: A FOCUS ON HYDROGEN PRODUCTION TECHNOLOGIES	Engineering innovation in energy infrastructure is needed to reduce harmful emissions and limit global warming. Hydrogen offers potential solutions if it can be produced economically, at scale, and with net zero CO2 emissions (H2NZ). The engineering challenges are considerable, and the recently launched Global Hydrogen Production Technologies Center (HyPT) aims to meet them. With funding from UK, Canadian, Australian and US Governments, HyPT focusses 27 organisations across 6 countries on advancing developments in three H2NZ technologies - water electrolysis, methane pyrolysis and photocatalytic solar water splitting. The technologies are considered together with three cross-cutting challenges across policy, water resources and markets. The cross-cutting challenges recognise that in transforming our energy infrastructure systems the engineering challenges are not just technical but bring about wider system change with justice implications. In developing solutions, engineering research must be cognisant of justice issues and the 'just transition' and take its share of responsibility in how technology is shaped. This poses real problems in practice in how to integrate technical and justice issues in a meaningful way in infrastructure engineering research. As part of the cross-cutting challenges, the project has outreach and engagement activities with indigenous peoples in US, Australia and Canada. This paper reflects on the journey to directly integrate engineering research on H2NZ production technologies, with justice and energy sovereignty implications for these communities. It discusses a matrix of questions, tailored to the different technologies, that seek to 'operationalise' and integrate justice principles for engineering research. The work has prompted reflection both on the role of engineer and researcher, and the meaning of energy justice for these communities. Can justice questions help support a reflective and responsible approach in engineering practice, re-shape problem-statements to embrace a wider-systems (yet contextualised) view, and better articulate the trade-offs that follow infrastructure adaptation? The aim is to initiate a transdisciplinary road map to explore how justice and sovereignty issues can speak to engineering practice on the ground. This paper reflects on the process to date leading to the development of the question matrix, challenges, and how it will be further developed. It explores the extent justice concepts, developed from other disciplines, can be integrated into engineering research in a meaningful way. In doing so it reflects on how justice questions, could and should inform engineering best practice. The paper is formed in three sections: 1.Academic grounding. Energy justice conceptualisation for the purpose of the paper draws upon environmental justice and capabilities approaches, before exploring how the justice literature integrates with energy sovereignty for indigenous communities. 2. Engineering. The technical differences between the three hydrogen production technologies and their societal and justice impacts are identified, noting differences in land and freshwater needs, the energy produced and use, waste and recycling potential, and the need for rare trace metals. 3. Research Support. Integrating the findings from 1 and 2, the matrix of questions is discussed. Its role (and limitations) in gathering a wider systems view of differing hydrogen production technologies and justice impacts are explored. Acknowledgement The authors gratefully acknowledge the financial support of the UK Engineering and Physical Sciences Research Council (EPSRC) for the Project entitled 'Global Hydrogen Production Technologies (HyPT) Center' under grant number EP/Y026098/1.

Joost Alleblas	**12A**	Rethinking Sufficiency	Rethinking Sufficiency Abstract Economics is often understood as the study of scarcity (i.e. of goods and resources). Abundance has been less examined, despite early explorations by Veblen and Galbraith. More and more, however, socio-economic problems of abundance have moved to the foreground of analysis. These problems concern the efficiency of markets, physical health, and waste. Furthermore, in between the extremes of abundance and scarcity we find sufficiency, a central notion in political economy. In relation to these three economic concepts, this study identifies three moral pairs of virtues and vices: generosity/excess (ethics of abundance), austerity/lack (ethics of scarcity), and moderation/stagnation (ethics of sufficiency). Of these virtues, moderation stands central to moral prescriptions of individual consumption while facing socio-economic abundance or excess. At the same time, moderation prepares us for periods of scarcity, as well as possibly preventing the manifestation of lack. Moderation is, finally, a central virtue in many religious and philosophical traditions. However, this study argues, such an account of moderation as individual virtue is problematic in the context of renewable energy, especially the generation of solar and wind power. First, moderation itself has connotations of stagnation and mediocrity, as moderation does not acknowledge the creative potential of situations characterized by energy lack or excess. Second, in order to create sufficiency, renewable energy inevitably suffers from excessive production, giving the intermittent nature of renewable sources such as wind and solar. Therefore, on the production side of these renewables, efforts to guarantee sufficiency, easily lead to energy abundance or morally problematic situations of excess, in which energy has to be squandered and costs incurred—even if technological solutions are provided. Third, the costs of wasting energy are not always fairly distributed. Fourth, as individuals have become prosumers of solar and wind energy, they themselves are confronted with situations of excessive production for which moderate consumption is an inadequate response. This study discusses several strategies for dealing with these issues, and rethinking the ethics of sufficiency in the context of renewable energy. Central to these strategies is the idea that other, collective virtues need to come into play to avoid morally problematic situations of excess. With this discussion, this study contributes to the reassessment of virtues important for the transition towards renewable energies. Keywords: Renewable energy, Civic virtues, Excess, Moderation, Prosumers
Janine Gondolf, Stefanie Enderle, Sophie Kuppler	**12A**	Contested Engagement: A Conceptual Analysis of "the Stakeholders" in Large Infrastructure Projects	In the face of global challenges, politics, science, and civil society must work together to develop transformation programs of varying focus and scope. Transformation programs, e.g. in response to crises such as the energy or digital transitions, create dramatic breaks with the actual state of social coexistence. These processes require a restructuring of the fundamental socio-technical constellations in society. They therefore have a potential impact on all social subsystems, making them complex and potentially conflict-prone. Stakeholder engagement activities of various kinds are prominent among the various methods used to open debate, exchange and participate. However, "stakeholder" is not a single concept, but rather a variety of methods, implications, research agendas, and disciplinary backgrounds (de Bussy and Kelly 2010, Weingart et al 2021). These implicit interrelationships of demands and expectations have received little conceptual attention, even though conceptualizing the role and scope of stakeholders is relevant to the successful engagement and impact of such activities. Especially in technology-based societies, people expect, at least in part, that new technological innovations will solve their current and future problems. Therefore, technology assessment (TA) (Grunwald 2019), as a fundamentally problem-oriented but also solution-oriented science and consulting practice, is actively involved in projects that involve stakeholders at all levels. Like other approaches in Science and Technology Studies (STS) (e.g. Konrad et al. 2016), TA's expertise lies in its interdisciplinary and integrative spectrum of theories, methods and processes for the critical examination, evaluation, communication and co-design of (evolving) technologies. The specific normative foundations of TA (democratic, sustainable, socially just and inclusive, issue-oriented) are indispensable for its inherent interdisciplinarity and transdisciplinary cooperation (with the public, politics, society). Specifically, TA is motivated by existing societal (not primarily scientific) problems and the creation of enabling structures that generate solutions in order to provide scientifically sound knowledge for negotiation and advice in political and societal decision-making processes. This contextualized view and practice of TA (e.g.; Böschen et al. 2021) opens up possibilities for more reflexive participatory action. In our paper, we outline a theoretical and methodological concept of "stakeholder" for multi- and transdisciplinary project activities and offer a nuanced assessment of the concept from our project work. It is intended to help both those who work with and those who study stakeholder engagement: Which stakeholder concept is helpful in determining with whom to engage in dialogue? What does dialog mean in a receptive context? How can responsibilities, processes, etc. be assessed and specified, since constellations may change during the process, e.g. formal responsibilities at the beginning may change over time? Keywords: stakeholder, participation, social change, technology assessment, decision-making processes References: Böschen, S.; Grunwald, A.; Krings, B.-J.; Rösch, C. (Eds.) (2021): Technikfolgenabschätzung : Handbuch für Wissenschaft und Praxis.Nomos. de Bussy, N.M. and Kelly, L. (2010), "Stakeholders, politics and power: Towards an understanding of stakeholder identification and salience in government", Journal of Communication Management, Vol. 14 No. 4, pp. 289-305. https://doi.org/10.1108/13632541011090419 Grunwald, A. (2019): Technology Assessment in Practice and Theory. Routledge Konrad, K.; van Lente, H.; Groves, C.; Selin, C. (2016): Performing and Governing the Future in Science and Technology. In: Felt, U.; Fouché, R.; Miller, C.; Smith-Doerr, L. (Eds.): The Handbook of Science and Technology Studies. 4. Aufl., Cambridge: MIT Press, 465–493. Weingart P, Joubert M, Connoway K (2021) Public engagement with science—Origins, motives and impact in academic literature and science policy. PLOS ONE 16(7): e0254201. https://doi.org/10.1371/journal.pone.0254201

Eva Pöll	**12B**	Engineering the Trust Machine	Complex technology has become an essential aspect of everyday life. We rely on technology as part of basic infrastructure and repeatedly for tasks throughout the day. Yet, in many cases the relation surpasses mere reliance and evolves to trust in technology. A new, disruptive technology is blockchain. It claims to introduce trustless relationships among its users, aiming to eliminate the need for trust altogether – even being described as "the trust machine". This paper presents a proposal to adjust the concept of trust in blockchain applications with the tools provided by conceptual engineering. Its goal is to propose a concept of trust, that offers more than a halo term, devoid of the normative depth that the original, philosophical term actually carries. As trust in blockchains connects information technology and philosophy, it needs both, those on the one hand, who understand the technological aspect of blockchain well, and those, on the other hand, who have expertise in ethics and philosophy. So far, however, the research in this area is rather thin. This research gap results in ambiguous terminology, overestimation of certain blockchain feautres, as well as an ambiguous understanding of trust in blockchains. One option to bridge this gap could be conceptual engineering. To use its tools and to find a suitable concept of trust in blockchain applications I suggest five criteria: A concept of trust in blockchain applications should (a) encompass and explain more than merely relying on technology; (b) reflect the functions of trust in technology: simplification and risk management; (c) be particularly fitting to the context of blockchain applications; (d) align with the necessary and jointly sufficient conditions for interpersonal trust or for trust in technology (condition of vulnerability, condition of expected competence, condition of willingness, condition of reliance); and (e) emphasize that the trust relation emanates from the trustor. These criteria ensure that the conceptual relation indeed embodies trust, thereby being richer than mere reliance, and that the chosen concept highlights the active role of the trustor. This also allows to evaluate possible concepts of trust in blockchain applications: While the concepts of trust in engineers and trust in algorithms have to be discarded, institutional trust emerges as a viable candidate, given some refinements. Ultimately, the concept of trust by normative expectations and attribution is suggested to capture the essence of trust in blockchain applications.
Ibo van de Poel	**12B**	Acceptable risk under moral uncertainty	I explore how to decide about acceptable (technological) risks under moral uncertainty. This question brings together two bodies of literature. One is risk ethics, which has addressed question about acceptable risks (and acceptable risk imposition), but has largely ignored issues of moral uncertainty (i.e., uncertainty about what ethical theory or moral principle is right). The other is the (recent) philosophical literature on moral uncertainty, which has by and large ignored descriptive uncertainty (and risks). Real-world cases almost always combine moral and descriptive uncertainty. Consider, for example, decision-making about lock-downs in the Covid-19 pandemic. Such decisions typically involve descriptive uncertainty about actual effects of decisions, like the effects on (public) health, freedom, well-being, mental health, etc. But they oftentimes also involve moral uncertainty about how important various moral principles (or values) are or about whether we should we should apply a utilitarian or a deontological ethical approach. Moral uncertainty is uncertainty about what moral theory to apply (MacAskill et al., 2020); or a bit more general: uncertainty about what normative framework to apply. Since different normative frameworks may order options differently, the question arises how to choose under moral uncertainty. MacAskill et al. (2020) propose maximizing expected choiceworthiness (MEC), a procedure that is largely analogous to maximizing expected utility for descriptive uncertainty. The question that arises is whether MEC can also be applied to questions about acceptable risks in which we have both descriptive and moral uncertainty. The standard form of MEC assumes descriptive certainty, or at least it assumes that different normative frameworks treat descriptive uncertainty in a similar way (i.e., similar to expected utility). However, that assumption seems false (Hansson 2003). Another problem of applying MEC to decisions about acceptable risk is that it frustrates the justification of decision about acceptable risks. MEC may allow a procedural justification ("the required decision-making procedure according to MEC was followed"), but it is unclear how the reasons that inform the normative frameworks that are considered add up or are combined. For example, if government had used MEC to decide about lock-downs, they best justification they would be able to give would be something like: "The calculations according to MEC show that a lock-down has the highest expected choiceworthiness." Such a justification would seem overly technocratic, and not very convincing and satisfying as a justification of such an important decision As an alternative I therefore explore two alternative procedures for deciding about (acceptable) risks under moral uncertainty. One is regret avoidance that aims to avoid the worst possible outcome in terms of moral evaluation. The other is coherence that frames the decision situation as one in which different normative frameworks may be partially right, rather than as one in which just one framework is right but we do are uncertain which one. MacAskill, William, Krister Bykvist, and Toby Ord. 2020. Moral uncertainty. New York: Oxford University Press. Hansson, Sven Ove. 2003. "Ethical criteria of risk acceptance." Erkenntnis 59:291-309.

Daniel Marom	**12D**	The contribution of theology to the philosophy of engineering: the example of the story of the Tower of Babel	The Contribution of Theology to the Philosophy of Engineering: The Example of the Story of the Tower of Babel The recent 2022 republication of the groundbreaking 1984 two-volume anthology Theology and Technology, invites those invested in the philosophy of engineering to reconsider the unique contribution that theological thought can make to the field. Like other areas of humanities study, theology adds a ripe reference point for the kind of reflexive understanding Carl Mitcham claims is necessary for engineers in the essay that dons the f-PET website, "The True Grand Challenge for Engineering: Self-Knowledge." One needs to put engineering on a bigger canvas to be able to truly see it. Theology effectively stretches that canvas to the point at which one can reflect on technological development in relation to the larger purposes of human existence, the limitations of human knowledge, the psychology of power, and the dangers in uncompromising pursuit of control over nature. Theology's potential contribution should also be considered in view of the less dichotomous approach that Eastern philosophy of engineering has taken to local religious and spiritual traditions such as Confucianism, Daoism and Buddhism (Miller et al, 2020; Wang, 2018; Zhu, 2010). Continuities and conflicts between spiritual traditions and modern approaches to science and technology can be highly valuable for the education of engineers and citizens alike. To explore possibilities in the context of Biblical religions, the paper offers an analysis of the Tower of Babel story from chapter 11 of the Book of Genesis, a theological tour de force that is surprisingly unaddressed in the republished anthology despite its emphasis on Christian exegesis. Popular perception is that this story is about humans' defiant rebellion against God culminating in their resounding defeat. Close reading reveals that it is in fact a sophisticated critique of total faith in technology as enacted by a pre-modern equivalent of a global community of engineers in contradistinction to God's mode of creation. This critique focuses on the denial of the extra-material realm of human existence and the infinite mystery of the universe in which it is embedded. The crux is in the ironic implosion of the whole project of the building of a city and a tower reaching to heaven. Rather than achieving their aim of "making themselves a name," – i.e. attaining full control over their fate - the builders experience a confusion of tongues that engenders their dispersal from Babel with the unfinished tower still standing. A second irony is in the chiastic tower-like structure and poetic technical mastery of the wording of the story. Examples will be presented of how these themes were explored through centuries of exegetical and artistic interpretation of the story – including recently, by the leading social psychologist Jonathan Haidt in application to the political effects of social media (Haidt, 2022). The paper then concludes with a discussion of the contemporary implications of the story in relation to - 1) the globalist aspirations of technological development, 2) technology's contribution to the dissolution of human attention and 3) the spiritual potentiality of engineering drawn from a comparison with the Burning Man event that is popular with hi-tech engineers. --- Mitcham, C., Grote J., Checketts L. eds. (2022), Theology and Technology: Volume One: Essays in Christian Analysis; Volume Two: Essays in Christian Exegesis and Historical Theology (Eugene, Oregon: Wipf and Stock). Miller, Glen, Wang, Xiaoweit (Tom), Sethy, Satya Sundar & Atsushi, Fujiki, eds. (2020) "Eastern Philosophical Approaches and Engineering" in The Routledge Handbook of the Philosophy of Engineering (New York: Routledge) Wang, Nan (2018), "Ancient Chinese Attitudes Towards Technics: Chinese Philosophy of Technology Prior to the 1800's," in Philosophy of Engineering, East and West (New York: Springer), pp. 61-73. Qin, Zhu (2010), "Engineering ethics studies in China: Dialogue between traditionalism and modernism," Engineering Studies, 2:2, pp. 85-107. Haidt, Jonathan (2022), "Why the Past 10 Years of American Life Have Been Uniquely Stupid" in The Atlantic, April 11.
Muhammad Abubakr	**12D**	Are Engineered River Basins Sentient? A Perspective from Islamic Philosophy	The idea that nature is a sentient being or possesses sacred qualities does not find much resonance with modern Western constructs. Yet, in recent years legal personhood has been granted to some rivers in India and New Zealand to safeguard them against excessive pollution and other damages (Edirisinghe et al. 2024). The law in New Zealand reflects and acknowledges the aboriginal worldview of the Whanganui Iwi tribe while in India, the evolving status of the Ganges and Yamuna rivers is driven by Hindu Vedic ontologies (O'Donnell and Talbot-Jones 2018). In the Muslim world, nature in general and rivers, in particular, are neither revered as deities nor objectified as completely disenchanted and mechanistic. Initiatives such as the 'Living Indus' in Pakistan (Butt 2023), and the popular perceptions of the 'dying rivers' among Muslim populations (Roberts 2005) around the globe indicate that a form of life is indeed attributed by Muslim cultures to inanimate yet highly complex natural processes such as rivers. This provokes us to ask the question: How are life and the self defined in the Islamic worldview? In this paper, we perform an inquiry into whether sentience or self-consciousness is a valid attribute for rivers from an Islamic philosophical perspective. We expand the definition of the river by coupling the natural processes with social processes to get engineered river basins, in which technologies such as irrigation and floodplain modifications play a transformative role in changing the river processes, and cybernetic technologies (such as AI, IoTs, satellite remote sensing) provide the key elements for information flows, feedback, and self-regulation within the engineered basin. Such an inquiry resonates strongly with the recently concluded scientific decade of the International Association of Hydrological Sciences (IAHS) promoting cross-cutting studies on dynamically coupled human-water systems (Montanari 2013). Our approach draws comparisons and inspirations from the process philosophy of nature championed by philosophers like Alfred N. Whitehead (1929) and various perspectives on the biology of life and mind that juxtapose being with becoming (Heimann 1942). Prominent scientists such as Ilya Prigogine (1982), James Lovelock (1979), and Fritjof Capra (2014) have used dissipative thermodynamics in biology, earth system physics, and ecosystems to expand on the definitions of life, cognition, and consciousness (Maturana and Varela 1991). In the Islamic philosophical tradition, parallels to such dynamic perspectives of natural processes can be drawn from the works of Mullah Sadra (Heriyanto 2021), Iqbal (Whittemore 1956), Rumi, and Ibn Arabi (Alpyagil 2007). Classical literature in Sufism and Islamic theology portrays the Self or a fully evolved Soul associated with the human individual (Homo Sapiens) as a complex evolving process composed of the Ruh (Spirit), Qalb (Heart), 'Aql (Mind), and Nafs (Ego) undergoing various states of dominance and suppression by the four constituents (Rothman and Coyle 2018). If a river basin is sentient, what will a model of the self for a river basin look like? What elements might be missing? The paper attempts to bring together these perspectives to look at river systems of the 21st century which have been massively instrumented and physically modified in the recent past and are currently at various levels of sustainability. We, therefore, consider natural (river) systems that have significant cybernetic elements (Wiener 1948) or embedded machine intelligence and which invoke the consideration of cyborg beings (Swyngedouw 1996). This provides additional elements to conceptualize a possibly sentient being, even without some of the key elements found in more evolved beings. The paper concludes with a discussion on whether the attribution of life and sentience changes the perspective of the basin managers who draw their worldview from Islamic cultural and religious traditions (Wescoat and White 2003; Wescoat and Muhammad 2021). In 2016, at a conference on Norbert Weiner in the 21st Century, the author reflected on the "onset" of cybernetic revolutions in complex river basins such as the Indus (Muhammad 2016). In this paper, we speculate on what might be the "culmination" of introducing smart information-driven technologies at a large scale in such river basins of the Muslim world. Bibliography Edirisinghe, Asanka, and Sandie Suchet-Pearson. "Nature as a sentient being: Can rivers be legal persons?." Review of European, Comparative & International Environmental Law (2024). O'Donnell, Erin L., and Julia Talbot-Jones. "Creating legal rights for rivers." Ecology and Society 23, no. 1 (2018). Butt, Mominyar Khalid. "Living Indus initiative of Pakistan: a major step towards climate change crises." Global Village Space (2023). Roberts, Tom. "The Indus–life-blood of Pakistan." Asian Affairs 36, no. 1 (2005): 1-11. Montanari, Alberto, Gordon Young, H. H. G. Savenije, Denis Hughes, Thibaut Wagener, Li Liang Ren, Demetris Koutsoyiannis et al. ""Panta Rhei—everything flows": change in hydrology and society—the IAHS scientific decade

Jesse Pappas	**12E**	Character Ethics Education for a Metamodern World	Engineering education faces the significant challenge of preparing students for professional success and ethical responsibility in an increasingly fluid world characterized by plurality and the potential liquification of character traits. Standard bearing approaches to character ethics education continue to target the gradual, trait-level change leading to behavioral consistency. Designed for the bygone modern era, this pedagogy is increasingly irrelevant to situatedness that now characterizes student identities. This paper lays a flexible foundation for a novel framework grounded Aristotelian Virtue Ethics and adapted to the emerging realities of contemporary society. Complementing traditional methods with late/metamodern strategies, this framework highlights ethical presence, adaptability, resilience, and motivation. The paper begins by addressing challenges inherent to liquid modernity and social acceleration, as conceptualized by Zygmunt Bauman and Hartmut Rosa, respectively. These philosophies underscore the need for education that is responsive to dynamic and often precarious conditions of modern life. A prospective framework integrates practical wisdom (phronesis), golden mean virtues, human flourishing (eudaimonia), and shared value creation. This model is sensitive to situational and contextual challenges, helping students navigate complex ethical dilemmas with a clarity of purpose. Core aspects include durable learning teams, experiential learning, narrative pedagogy, near-peer mentorship, and community engagement. The framework also emphasizes the importance of mutually supportive social environments and cultivation of reflexivity and resilience. A brief exposition of implementation strategies and emergent challenges currently engaged at the University of Virginia precedes concluding remarks urging universities to reimagine engineering ethics education at scale and to consider adopting dynamic and adaptive pedagogies in non-technical learning domains.
Diana Martin	**12E**	Teaching engineering risk and uncertainty as ethical concepts	An important part of engineering practice requires decision-making in ill-structured contexts that are rife with risks and uncertainty (Jonassen, 2009). Such decisions have an ethical core since they have the potential to affect the safety and welfare of a wide range of stakeholders. To expose engineering students to ethical decision-making, a popular teaching method are case studies (Martin et al, 2021). Cases are typically focused on students' responses to hypothetical, obvious and brief ethical dilemmas or vignettes, bracketing the group dimension or real features of the context of engineering practice (Hoffmann & Borenstein, 2014; Bairaktarova & Woodcock, 2017). This approach has been criticized as simplistic, as is often framed from the perspective of decisions made in absolute knowledge or with the knowledge of the devastating aftermath of decisions, as in the case of tragedies such as The Challenger explosion. Another criticism is their emphasis on the perspective of the individual making the decision, bracketing the epistemic stance of other stakeholders. It is thus essential to introduce engineering students to ethical decision-making as situated practice, in a manner that can account for the complexities of real-life decision-making under risk and uncertainty. This contribution argues that engineering ethics education would benefit from a concerted focus on introducing to engineering students the concepts of risks and uncertainty as ethical concepts. It offers an example set in a Dutch Technological University, describing how these concepts were introduced in the course on Decisions Under Risk and Uncertainty offered by a Dutch Technological University in the academic year 2022/23. The teaching methods comprise live case studies brought by 4 guest speakers to the course who describe their own decision-making in face of ethical dilemmas comprising risk and uncertainty, and tools they employed such as community engagement and participatory tech assessment. The live case studies were brought by Zach Pirtle (NASA), Ben Pauli (author Flint Fights Back), Laura Nolan (whistleblower Google Project Maven and volunteer for Stop Killer Robots) and Diana Bairaktarova (Virginia Tech). The contribution makes the case for rethinking engineering ethics case studies to incorporate epistemic dimensions related to risk and uncertainty, as well as the implications of this focus in terms of ethics learning gains, that contributed to making students question epistemic expertise and becoming aware of the manifestation of power relations in engineering decision-making. References Bairaktarova, D. & Woodcock, A. (2017). Engineering students' ethical awareness and behavior: A new motivational model. Science and Engineering Ethics, 23(4), 1129-1157. Hoffmann, M. & Borenstein, J. (2014). Understanding ill-structured engineering ethics problems through a collaborative learning and argument visualization approach. Science and engineering ethics, 20(1), 261-276. Jonassen, D.H., Shen, D., Marra, R.M., Cho, Y.-H., Lo, J.L., & Lohani, V.K. (2009). Engaging and supporting problem solving in engineering ethics. Journal of Engineering Education, 98(3), 235–254. Martin, D.A., Conlon, E. & Bowe, B. (2021). Using case studies in engineering ethics education: the case for immersive scenarios through stakeholder engagement and real-life data. Australasian Journal of Engineering Education, 26(1), 47-63.
Mareike Smolka, Bas Boom	**12F**	Imagining AI in Context: (Re-)analysis of an Interdisciplinary Student Project	Evocations of utopian and dystopian futures proliferate in policy, corporate, and media discourses on Artificial Intelligence (AI), where AI is presented as the ultimate fix to global challenges and as a threat to humanity. In particular, the agricultural sector is pervaded by hype/fear imaginaries: on the one hand, the European Commission designates it as "high-impact sector" where smart robots can improve farm efficiency and environmental sustainability; on the other hand, farmers fear that AI-enabled precision farming will increase the power of large corporations at the expense of small businesses and create considerable environmental footprints throughout technological lifecycles. Dominant hype/fear imaginaries leave little room for imagining alternatives because the future with AI is portrayed as fixed and inevitable. This reflects a crisis of imagination: simplistic technology-driven futures are easy to imagine, but it is difficult to anticipate plausible ways in which technology and society could be interrelated within future socio-technical systems. AI researchers and developers have sought support from social scientists and humanities scholars to engage with 'alternative visions', 'bottom-up imaginaries', and 'counter-narratives'. Their collaborations have drawn on visual art, science-fiction writing, indigenous knowledge, social theory, stakeholder and public engagement. Stakeholder and public engagement does not only promise to align technology-driven futures with society-oriented alternatives, it also renders AI research and development more legitimate. By involving stakeholders and publics as co-designers in research and development workflows, those creating novel AI technologies become accountable to those affected. Building on engagement approaches at the intersection of Science & Technology Studies (STS), Philosophy of Technology, and Design Studies, we have developed the "Imagining AI in Context" student project. Under the supervision of an STS scholar (Mareike Smolka) and a computer vision researcher (Bas Boom), students from social, natural and engineering disciplines conduct a six-week project in which they initiate a co-design process to imagine smart robot designs for precision pruning in orchards. The project is situated at OnePlanet Research Center, a high-tech innovation center connected to Wageningen University, where AI-powered technologies for precision health and agriculture are developed. Students are tasked to develop robot designs while attending to the following questions: How do stakeholders and publics imagine smart robots for pruning, thinning, spraying, and other tasks in orchards? What are their concerns regarding economic viability, environmental sustainability, social justice, and other values? How could a business case look like that takes these values into account? In this talk, we present the analysis and results of the student project. Based on interviews with stakeholders involved in fruit growing, expert pruning, and orchard machinery, students identified an imaginary of automation. However, a re-analysis of the data – paying attention to alterity by reading the interviews against the grain – revealed tensions with and alternatives to this imaginary. We conclude that the future with AI may not be as fixed as it appears. By keeping unpacking and opening tensions in otherwise coherent accounts, we recognize possibilities for aligning a technology-driven imaginary with alternative futures that may be desirable from a social, environmental, and political perspective.

SaeBom Song, Andreas Lösch	**12F**	Openness as Smartness?: Sociotechnical Evolution and Imaginaries of Open Data Initiative and Smart City in the Republic of Korea	This paper argues for approaching the open data initiative as a socio-technical imaginary within the evolving socio-technical narrative of the smart city in Korea. By conducting a discourse analysis of primary source material from prominent public websites and official public documents, forums, and social media over a decade, as well as the 11 narrative interviews, I argue that the socio-technical imaginaries of openness and smartness in Korea have evolved in their different and pluralistic path. However, amidst the global embrace of open data and smart city, I want to identify how the globally shared values and expectations of open data and smart city are being shaped, evolved, and developed in the nation-specific context of South Korea. In the last decade, Korea is well known for technical innovations and the rapid implementation of the vision of a smart city. Korea is a significant example of how smart city policy has evolved over time, demonstrating that smart city policy is changing dynamically. In the early stages of smart city development in Korea, the national focus is on top-down techno-driven adoption to solve urban challenges by actively establishing digital infrastructures and services. The smart city driven by digital technologies was the tangible manifestation of an envisioned political goal and the projection of smartness onto the city as a whole, dominated by a global corporate discourse. Beyond this technological determinism and corporate narrative, sociotechnical narratives of the smart city in Korea have recently underlined that the autonomous participation and collaboration of civil actors will foster greater sustainability and resilience in urban planning in the long run. Apart from smart city initiatives. the open data initiative has become one of the most successful digital infrastructures since 2012. In the beginning, the implementation of open data initiatives was strongly driven by distinctive national digital policies, a vigorous commitment to the administrative transformation of local government, and local-specific social issues. However, the concepts of smart city and open data initiative have ultimately become complementary, and each can leverage the other's narrative to reinforce the development logic of the respective projects. This complementary relationship between smart city and open data has been successfully intertwined over the past few years, accelerating the formation of integrated policies and their performative realization in the socio-technical imaginaries of both. Furthermore, I argue that the premise of the smart city and open data imaginary of the Republic of Korea is strongly rooted in a particular narrative that emphasizes technological solutions to achieve more economic growth, political participation, better urban management, and the resolution of urban crises. Such a sociotechnical framing of shared beliefs, values, and expectations surrounding the smart city and open data initiative in Korea creates a visionary and reactionary force and presents itself as “politically benign and commonsensical (Kitchin, Lauriault and McArdle, 2015)”. However, the Seoul Metropolitan Government's open data initiative reveals that there is a discrepancy between the actual use of open data by citizens and the private sector and the sociotechnical narratives pursued by open data policies. Consequently, this work addresses the limitations to the sociotechnical imaginary of open data initiatives in Korea and suggests that we require to reimagine what openness and smartness can mean and create counter-narratives that open up space for alternative values and models. Reference: Kitchin, R., Lauriault, T.P. and McArdle, G. (2015) ‘Smart cities and the politics of urban data’, in Smart Urbanism: Utopian Vision or False Dawn? London: Routledge, pp. 16–33. Available at: https://mural.maynoothuniversity.ie/7323/ (Accessed: 4 January 2023). Keywords: Smart city, Open data, Sociotechnical imaginary, Sociotechnical systems, STS, Smart technology
Paula Gürtler, Artur Bogucki, Berta Mizsei	**12F**	Ethical, Legal, and Socioeconomic Aspects in AI projects	Risk management of Artificial Intelligence (AI) systems has emerged as a primary concern in an era marked by rapid developments in the field. The social sciences and philosophy play a crucial role in assessing both the ex-ante and ex-post ethical, legal, and socioeconomic impacts of the technology. For European-funded projects, the methodological framework of Ethical, Legal, and Social Aspects (ELSA) has recently gained renewed attention for the evaluation of these initiatives (Ryan & Blok, 2023). Although there is a considerable body of literature discussing ELSA (Ryan & Blok, 2023; Zwart et al., 2014), guidance for researchers seeking to apply this method in practice remains scarce. With the AI Act coming into force, there is an increased need to develop methodologies for the two primary enforcement mechanisms of the Regulation: conformity assessments and post-market monitoring of algorithmic systems. Building on the ELSA methodology promises a comprehensive approach to addressing these requirements with respect to sectoral auditing of safety legislations. The objective of this paper is to contextualise this research within the ELSA history, explain why we propose to adopt an ELSE (ethical, legal, socioeconomic) aspects framework instead, and to provide detailed and concrete methodological guidance, akin to a manual. Relying on existing literature on this topic (Ryan & Blok, 2023), a list of core principles can be identified for the ELSA methodology. To avoid reinventing the wheel by transitioning from RRI back to ELSA, we propose a synthesis of these two approaches. This involves broadening the scope of ELSA to include economic aspects, thereby transitioning to the ELSE framework. This change means taking on board the lesson learnt from RRI on the importance of economic sustainability of (publicly) funded AI development. The lack of practical guidance for conducting ELSE aspect analysis limits the potential usefulness of this framework. Existing examples of ELSA research either do not reflect the participatory and stakeholder engagement that ELSA promises on paper (Adhikari et al., 2024; Chandler et al., 2022) or do not cover the full range of disciplinary aspects (Stahl & Leach, 2023). We aim to expand the approach by offering practical and concrete methodological guidance on implementing ELSE in research projects. ELSE incorporates a variety of methods, with this discussion focusing specifically on stakeholder involvement. This includes methods for stakeholder identification and the design of focus groups. The goal is to share lessons learned from utilizing the methodology in research projects and provide the practical guidance that is missing from the current literature. References Adhikari, K., Naik, N., Hameed, B. Z., Raghunath, S. K., & Somani, B. K. (2024). Exploring the Ethical, Legal, and Social Implications of ChatGPT in Urology. Current Urology Reports, 25(1), 1–8. https://doi.org/10.1007/s11934-023-01185-2 Chandler, J. A., Van Der Loos, K. I., Boehnke, S., Beaudry, J. S., Buchman, D. Z., & Illes, J. (2022). Brain Computer Interfaces and Communication Disabilities: Ethical, Legal, and Social Aspects of Decoding Speech From the Brain. Frontiers in Human Neuroscience, 16, 841035. https://doi.org/10.3389/fnhum.2022.841035 Ryan, M., & Blok, V. (2023). Stop re-inventing the wheel: Or how ELSA and RRI can align. Journal of Responsible Innovation, 10(1), 2196151. https://doi.org/10.1080/23299460.2023.2196151 Stahl, B. C., & Leach, T. (2023). Assessing the ethical and social concerns of artificial intelligence in neuroinformatics research: An empirical test of the European Union Assessment List for Trustworthy AI (ALTAI). AI and Ethics, 3(3), 745–767. https://doi.org/10.1007/s43681-022-00201-4 Zwart, H., Landeweerd, L., & Van Rooij, A. (2014). Adapt or perish? Assessing the recent shift in the European research funding arena from ‘ELSA’ to ‘RRI’. Life Sciences, Society and Policy, 10(1), 11. https://doi.org/10.1186/s40504-014-0011-x

Matthias Razum, Anna Jacyszyn, Linda Nierling, Felix Bach	**12F**	Leibniz ScienceCampus „Digital Transformation of Research“	The digital transformation of research affects both academia and society as such. It comprises the adaptation of epistemological processes as well as the transparency and communication of findings. New data-driven methods not only open up innovative approaches for knowledge generation, but also raise legal, ethical and societal questions. Within the newly established Leibniz ScienceCampus “Digital Transformation of Research” (DiTraRe) we investigate the effects of the increasing digitalisation of scientific work in four research clusters, in which interdisciplinary teams of researchers, computer scientists, lawyers and sociologists work together.Based on specific challenges (use cases), the research clusters aim to develop practical solutions. In the next step, we try to apply these solutions to further use cases and thus generalise them iteratively. In the cluster “Protected data spaces”, we analyse the use of sensitive data in sports science.Thecluster “Smart data acquisition” deals with Smart Lab technology such as Laboratory Information Management Systems (LIMS) and Electronic Lab Notebooks likeChemotion.Thecluster “AI-based knowledge realms” focusses on artificial intelligence and machine learning in biomedical simulation and engineering. The topic of “Publication cultures” are changes in scientific publishing due to new challenges like the publication of large datasets with a use case originating in climate research. In addition to the research clusters, the ScienceCampus includes four dimensions that consider cross-cutting topics on a meta-level, for example in the areas of legal and ethical challenges, security, tools or resonance and reflection. The last dimension in particular expands the view from science to society and how society perceives the changes in science. This concerns trust and credibility of scientific findings as well as citizen science and the assessment of the value of research for society. In our contribution, we will present the concept of a Leibniz ScienceCampus and the specific characteristics of DiTraRe. One focus will be the interface between natural sciences and engineering on the one hand and sociology, technology assessment, law and ethics on the other. We will report on our experiences with the chosen interdisciplinary and multi-layered approach and present the first results of our work with a focus on current challenges, namely compare current practices of publication, organisation and data management in different research disciplines, assess current and future risks, e.g. in the field of data privacy and security, anticipate and assess changes due to AI, e.g. the use and impact of LLMs, and design of transdisciplinary exchange on research data using innovative science communication formats towards citizens and science-society intermediaries, e.g. science journalists.
Ritesh Bansal	**13A**	Distinguishing between Nudging and Usability in the context of Product Designs	Keywords-, Nudge, Usability, Design Innovation, Cognitive Biases, Libertarian Paternalism Abstract: We use many products and services every day, resulting from several design modifications after years of research by the companies’ engineers and innumerable user feedback. A new product is launched in the market after rigorous testing and modifications accordingly. Designers make modifications and innovations after testing and getting feedback from users to enhance usability. For example, we see continuous improvement and updates in computer hardware and software designs. Similarly, Wright brothers made history on 17 December 1903 with their first flight, which was damaged on the same day after flying a maximum of 59 seconds in the air. Nevertheless, that was not the end but the beginning of a new era. Now we have colossal aircraft flying around the earth and into space. Similarly, we humans make several decisions in our everyday life. We face choices where some options are better than others and our selection may lead us to different paths of our life. For example, we all know that eating fruits is good for our health but still following these good habits become hard in the real world when we have both fruits and cakes in front of us.Required rational human behavior is not always spontaneous. To overcome such lag between required behavior and actual behavior, there is a need of gentle push which may lead us towards the required behavior. Thaler & Sunstein in their seminal book titled ‘Nudge’ (2008) proposed that “the same principle of good design and functional architecture apply in the world of choices as well”. They introduced the concept “Nudge” according to which, it is possible to influence people’s behavior in a predictable way if the choices are designed by carefully considering human behavioral insights including various heuristics and biases. For example - Carolyn is in charge of the cafeterias of many schools in the city where thousands of kids eat. If she decides to keep the healthy foods such as fruits at eye level which makes them easy to get, while keep away unhealthy items away, it can increase the consumption of healthy foods and decrease the consumption of unhealthy ones. However sometimes while giving analogy between product designs and nudges, product automations such as automatic switch for the headlights are referred as nudges. They say “Any remotely sensible car comes with an automatic switch for the headlights that turns them on when you are operating the car and off when you are not, eliminating the possibility of leaving your lights on overnight and draining the battery. Nudges of these kinds are saving lives, and we can expect many more of them in the future.” Although these examples may resemble instances of nudging, they may not necessarily be classified as nudges. These instances are kind of product innovation where the focus is on the usability of the products. In this article, we will attempt to draw a distinction between nudges and usability measures by outlining the necessary conditions of nudges and usability measures from their definitions. The significance of making this distinction is that the ethical implications of the design and employment of nudges can be analysed in a more effective manner if we can distinguish what a nudge is and what it is not. Considering usability measures and similar automation as "nudges" can lead to ambiguity while analysing the moral consequences of the design and applications of nudges, and hence, making this distinction is inevitable to constitute productive, ethical discourses to critically analyse the design and application of nudges.
Clément Lasselin	**13A**	Engineering Research as a Science: Perspectives from the Philosophy of Science	Engineering research is sometimes distinguished from scientific research. This distinction stems from the fact that scientific research is often associated with characteristics that are hardly associated with engineering research. On the one hand, scientific research is often presented as a set of activities that lead to the production of the most accurate information possible about certain parts of reality, based on various criteria. On the other hand, engineering research is seen as a set of activities that lead to the production of information that favors the production of some goods or services. Engineering research does not have to produce the most empirically accurate information. Therefore, since scientific research and engineering research are not the same activities, they do not seem to be the same types of research. The purpose of this presentation is to clarify how engineering can be defined in relation to scientific research, and what the different implications of ways of defining engineering may be. The purpose of this presentation is also to explore other definitions of these two types of research. To do this, I will use elements of philosophy of science. Philosophy of science is the discipline that asks what are sciences, scientific information, scientific goals, or scientific methods in order to better distinguish scientific research from other types of research. Historically, philosophy of science has focused more on physics and biology, and less on engineering research as such. It seems that philosophers of science have rarely considered engineering research as a type of scientific research. Nevertheless, engineering research may be interesting to study for philosophy of science, at least as a kind of research on the borderline with proper scientific research. In particular, it may be helpful to consider the role of so-called epistemic and non-epistemic goals in the conduct of research. Using several examples in agricultural, energy, and chemical engineering from my dissertation on bioenergy, I will show that engineering research can be understood as scientific under certain conditions. I will argue that engineering consists in aiming to produce the most empirically accurate information while integrating other epistemic goals. I will show that this is common in scientific research, and that it is particularly important to consider how engineering research is conducted and for what purposes in order to consider whether or not it is scientific.

Miriam Vetter, Sonja Haug, Karsten Weber, Caroline Dotter	**13F**	Acceptance and Willingness to Use Smart Meter Applications. Results from a German Population Survey	Abstract: After industry (44%) and retail and services (27%), private households are the third biggest contributor to overall German electricity consumption in 2021 (Bundesverband für Energie und Wasserwirtschaft, 2022). Private household energy consumption is driven by sociodemographic, residential, and appliance-specific factors (Jones et al., 2015). Evidence from South Korea also suggests private energy consumption is driven by behavioral patterns, equipment use, and psychosocial aspects (Ryu & Kim, 2022). To achieve the transition to a carbon-neutral society, residential electricity consumption needs to be reduced and made more flexible reacting to peek demands in electricity supply. This is expected to be achieved partly by employing so-called smart meters. Smart meters are currently rolled out in Germany for private households. In our paper, we present the results of a national population survey in Germany conducted in November 2023 (n=2,027). The survey asked about awareness, acceptance, and willingness to use smart meters as well as application scenarios utilizing AI in private households. Half of the respondents were familiar with the term smart meter, while one-fifth stated current usage of a smart meter. While the majority of households consider smart elecricity meters as useful and understand how they could benefit from usage, expectations regarding usefulness and benefits are more pronounced in households that already have smart meters. Most respondents cite financial reasons for using smart meters; they hope to get cheaper energy and reduce their energy consumption. Consumers were presented with two scenarios of smart meter application: In scenario 1 an app displays real-time electricity consumption and identifies the devices consuming the most electrity. Scenario 2 extends this by additionally presenting consumers with energy saving tips. Scenario 1 receives higher approval compared to scenario 2. Consumers are interested about their electricity consumption patterns and the consumption of specific devives. However, they apparently do not want to receive suggestions to change their behaviour, but want to draw their own conclusions.
Arian Mahzouni	**13F**	The role of electric vehicle batteries in shaping multi-system interactions in urban energy systems: a socio-technical approach	As part of EMPOWER project, we have recently implemented a literature review. We found a range of interactions between EV batteries and other urban systems and technologies to save energy and to reduce environmental pollution, transforming the way energy services are delivered. The key forms of interactions are grid to vehicle (G2V) for charging EV batteries, vehicle to grid (V2G) for discharging EV batteries, and vehicle to building (V2B) as a substitute for grid system for powering buildings and for mitigating fluctuations from renewable energies (cf. Huang et al. 2022; Wang et al. 2023). Emerging resource flows among these systems calls for establishing new structural couplings to support associated activities of EV batteries, e.g., charging infrastructure and technology and battery regulations to enhance circular business models of EV batteries (cf. Baars et al. 2023). We will discuss if and how changes in structural couplings might create an enabling environment for new resource flows by drawing on the notion of 'duality of structure' (cf. Giddens 1981; Giddens 1984) and on the ontology of socio-technical structures about the need to include non-human material objects into the social structures (cf. Elder-Vass 2017a; Elder-Vass 2017b). The key research questions are: 1) How do actors and their associated activities generate new structural couplings between different urban systems related to EV batteries across the dimensions of actors, technologies, and institutions? 2) How can we combine the theory of 'duality of structure' with new approaches on socio-technical structures to better understand the key obstacles to effective interactions in urban energy systems? By undertaking an interdisciplinary approach, we will investigate how multiple system transitions interact with each other. This might help to unfolding many kinds of interactions in the form of both 'regime-regime' and 'regime-niche' constellations, e.g., among electricity supply system, EV batteries, grid system, EV charging system, PV and building system. References Baars, Joris, Felipe Cerdas, and Oliver Heidrich. 2023. 'An Integrated Model to Conduct Multi-Criteria Technology Assessments: The Case of Electric Vehicle Batteries'. Environmental Science & Technology 57 (12): 5056–67. https://doi.org/10.1021/acs.est.2c04080. Cao, Sunliang. 2019. 'The Impact of Electric Vehicles and Mobile Boundary Expansions on the Realization of Zero-Emission Office Buildings'. Applied Energy 251: 113347. https://doi.org/10.1016/j.apenergy.2019.113347. Elder-Vass, Dave. 2017a. 'Material Parts in Social Structures'. Journal of Social Ontology 3 (1): 89–105. https://doi.org/10.1515/jso-2015-0058. ———. 2017b. 'Materialising Social Ontology'. Cambridge Journal of Economics 41 (5): 1437–51. https://doi.org/10.1093/cje/bex038. Giddens, Anthony. 1981. A Contemporary Critique of Historical Materialism: Vol. 1 Power, Property and the State. London: Macmillan. ———. 1984. The Constitution of Society: Outline of the Theory of Structuration. Cambridge: Polity Press. Huang, Pei, Ran Tu, Xingxing Zhang, Mengjie Han, Yongjun Sun, Syed Asad Hussain, and Linfeng Zhang. 2022. 'Investigation of Electric Vehicle Smart Charging Characteristics on the Power Regulation Performance in Solar Powered Building Communities and Battery Degradation in Sweden'. Journal of Energy Storage 56: 105907. https://doi.org/10.1016/j.est.2022.105907. Wang, Gang, Yuechao Chao, and Jianqing Lin. 2023. 'Technical-Economic and Environment Benefit Analyses of a Novel Building Attached Photovoltaic System'. Environmental Science and Pollution Research 30 (45): 100660–74. https://doi.org/10.1007/s11356-023-29530-5.
Manuel Baumann, Marcel Weil, Jens Peters, Hüseyin Ersoy, Merve Erakca, Haruna Bismark	**13F**	Constructive Technology Assessment to support sustainable battery development	Sodium-ion batteries (SIB) are considered as a promising alternative to Lithium-ion batteries (LIB), regarding sustainability challenges as carbon footprint, criticality, social impact and other aspects. There are several SIB types being developed with a wide set of different materials. However, a major challenge in early technology readiness levels (TRLs), here regarding research of SIB, is to identify most promising and sustainable pathways for potential commercialization. This is a challenging task as there is a high degree of uncertainty in early TRLs due to the lack of robust data. A design and decision dilemma arising from the claim to achieve something as "better or more sustainable battery technology" (is to find the right shape targets (e.g., environmental aspects, social aspects, supply security) and how to characterize these. This phenomenon is often associated with the so-called Collingridge dilemma, which states that: in early technology development stages opportunities to steer are plentiful, but hard to choose from, while at later stages this is reversed. Constructive Technology assessment offers a possibility to tackle this dilemma by broadening the design phase of new technologies by feedback of technology assessment activities into the construction of technology, here new SiBs. By nature CTA has a transdisciplinary research orientation which allows to; 1) incorporate processes, methodologies, knowledge, and goal of stakeholders from and across academia as well as actors from outside academia, 2) to create solution-oriented and social robust knowledge which is transferable to scientific and societal practice. Hereby, a conceptual framework is presented that maps methods that can be incorporated into CTA to provide different spheres of information that can support a sustainability by design approach using quantitative and qualitative approaches. The talk will provide an overview of methodological challenges for sustainability assessment that arise from early to high technology development levels of post-lithium battery research in the frame of CTA. Furthermore, the presented approach can serve as a blueprint for various battery systems or even other technologies. It can be seen that energy density is one of the most important factors, which strongly influences results. Chemistries containing nickel, vanadium and cobalt have in tendency higher impact in a craddle to gate perspective.

2) Experimental Sessions, Panels and Art Exhibitions

Rafael Mestre, Ned Barker, Sergey Astakhov, Aníbal M. Astobiza, Maria Guix, Joana Burd, Matt Ryan	**3G EXP. SESSION**	Interdisciplinary Speculations for the Future of Biohybrid Robots	Biohybrid robotics is a field that promises to revolutionise the current robotics landscape by combining biological entities with artificial materials. Some examples in the literature have included muscle-based biological swimmers or crawlers, AI-designed 'xenobots' made of frog cells or cyborg beetles. The field is still in its infancy, but it already brings together interdisciplinary research on biology, materials science, nanotechnology, and engineering. There have not been, however, deeper philosophical or sociological explorations of this field, apart from human-centred discussions on transhumanism, cyborg art, or embryonic stem cell research in biomedicine. The social, ethical or climate implications of this field, which stems from a series of fragmented frontiers in engineering research, are still not completely clear. In this experimental workshop, we propose an interdisciplinary session focused on the theme of Understanding, Assessing, and Designing Responsible Futures for Biohybrid Robotics. Drawing on our mixed expertise in deliberative, artistic, ethnographic, and STS methods, we bring together experts in biohybrid engineering, philosophy, art, sociology, and public policy to speculate about the future of this technology from a variety of disciplines. The following experts have confirmed their attendance: Maria Guix: chemist and bioengineer, expert in biohybrid robotics. Aníbal M. Astobiza, philosopher and expert in ethics of emerging technologies. Sergey Astakhov/Ned Barker: sociologists, experts on biohybrid bodies and sociology of expertise. Joana Burd: arts-based researcher, specialising in visual arts, technology, and the senses. Matt Ryan: political scientists, expert on public policy and AI. The experimental workshop will be organised in two main sessions. In the first half, the invited panellists will give short talks setting the ground for discussions on the futures of biohybrid robotics from their different perspectives. After this, the room (panellists and attendees) will be divided into three groups to take on separate activities based on structure prompts around the notion of interdisciplinary speculations, addressing the conference theme: Understanding: a mapping exercise on what we need to understand to develop future biohybrid robots (e.g., fundamental biology, systems approach, interaction design, AI). Assessing: deliberative exercise focused on how we assess that biohybrid robots are developed for social good (avoiding biological, environmental, and existential risks, regulation, policies). Designing: rapid speculative design of four different future scenarios based on the 4Ps method (possible, plausible, probable, preferred). After the group deliberation, the remaining 15 min will be spent on a group discussion, in which a group representative will share the output of the work and have a short Q&A. The workshop will be chaired/facilitated by Rafael Mestre, expert on biohybrid robotics and responsible research of emerging technologies. With this workshop, we hope to enable speculative yet useful discussions on the futures of this early stage emerging technology, ensuring inter and transdisciplinary views by bringing together not only engineers and philosophers that are experts in the technology, but also artists, sociologists and public policy experts.
Erhardt Graeff, Guru Madhavan	**6G EXP. SESSION**	Civic-minded Engineers and Wicked Problems	In his 2024 book Wicked Problems, Guru Madhavan entangles civic-mindedness with engineering, contending that "Developing a civic consciousness to achieve democracy's goals will fundamentally require a system engineering approach, just as remedying the professional deficits of engineering will depend on civic participation" (p. xv). This interview and conversation between Guru Madhavan, Norman R. Augustine Senior Scholar and senior director of programs of the National Academy of Engineering, and Erhardt Graeff, Associate Professor of Social and Computer Science at Olin College of Engineering, will illuminate the past, present, and potential future of what Madhavan describes as "an engineering vision for civics and a civic vision for engineering" (p. xv). We will draw from the ideas of philosopher Peter Levine and political theorist and civic educator Harry Boyte to define what civics and civic professionalism in engineering mean. We will explore the nature of "wicked problems" and the distinct qualities of systems engineering as an approach to these problems and also to the wider realm of democratic practice. Relying on Graeff's original research on civic-mindedness in engineering, we will examine the ways engineering and engineering education have moved toward and away from civic conceptions of an engineer's roles and responsibilities over time. By the end, we will discuss opportunities and barriers for promoting civic-minded engineering amidst other efforts to enhance ethics, responsibility, and value-sensitivity in engineering.

David Goldberg, Katherine Goodman, Zachary Pirtle, Diane Michelfelder, Daniel Marom	**7G EXP. SESSION**	Ways to Get Philosophy of Engineering Taken Seriously by Engineers	Roundtable discussion: Ways to Get Philosophy of Engineering Taken Seriously by Engineers The aim of the roundtable discussion is to invite members of the f-PET community to grapple with an issue that is fundamental to its impact on engineering: how to seriously engage engineers with philosophical study of their profession. Without such study, engineers lack the ability to reflect deeply about their role in transforming how and why we live. It is therefore in f-PET community's interest to better articulate the value of philosophy or engineering to engineers and to develop effective practical means to bridge the divide that separates them. This is not a narrow or technical matter. It can be addressed at many levels. At the macro level it relates to educational policy in higher education, lifelong learning in industry, professional regulation by government and leadership and faculty education for the profession. At the micro level, it relates to the content, curriculum, and pedagogy of teaching prospective or practicing engineers – even with respect to a single lesson or teaching interaction. At all levels, resistances to bridging the divide are so deeply embedded in the overall system of engineering and engineering education that strategic thinking and planning are critical. To begin this process, it is useful to address fundamental questions, such as: Since engineers have been educated and have practiced with minimal or no philosophical input for some time, why is it now important for them to engage more intensely with philosophy? What specific contributions could philosophy make to engineering education and practice? How would you complete the sentence: "Philosophy is important to engineering education and practice as a…?" Can you give one instance of such importance? What pre-conditions are necessary for efforts to make philosophy become important for engineering education and practice? What kind of policies, institutions, initiatives, and communications of its importance would make philosophy more attractive to engineering students, practitioners, faculty instructors, mentors, and the public? The roundtable will begin with a short introduction to the topic by the discussion leader, Diane Michelfelder, as well as a quick initiation exercise with the audience. Then, four discussants will present their ideas on the topic: David Goldberg, Zachary Pirtle, Katherine Goodman and Daniel Marom. The audience will then be invited to contribute responses and alternative perspectives. After that, each presenter will be invited to make one final point and the discussion leader will conclude with a few points of general summary and continuation. Participant bios: Discussion Facilitator - Diane P. Michelfelder: Professor Emerita of Philosophy, Macalester College; Past President of the Society for Philosophy and Technology; Co-founder of fPET; Co-editor of The Routledge Handbook of the Philosophy of Engineering. David E. Goldberg: Civil engineer; Professor Emeritus, University of Illinois at Urbana-Champaign; AI researcher; President and Leadership coach of Three Joy Associates; Co-author of the The Whole New Engineer: The Coming Revolution in Engineering Education; Co-founder of f-PET; Consultant on change in higher education and engineering education. Zachary Pirtle: Systems Engineer, program executive and senior policy Advisor; George Washington University, lecturer on systems engineering; f-PET leader and website curator. Katherine Goodman: University of Colorado at Denver, Associate Professor of Engineering and Director of Center of Excellence in Teaching and Learning; leadership of initiatives in engineering education that include elements of faculty training, curriculum development and research. Daniel Marom: Member of the founding faculty of the Mandel School for Educational Leadership; Developer of educational leadership training in medicine, engineering, psychotherapy, and other professions; Research Fellow at the Technion Israel Institute of Technology.
Filippo Santoni de Sio, Jordi Viader Guerrero, Aarón Moreno Inglés, Andrea Gammon	**8G EXP. SESSION**	Innovating educational approaches in engineering ethics: Experiencing space, power, and social relations, and Imagining responsible (technological) futures	Innovating educational approaches in engineering ethics: Experiencing space, power, and social relations, and Imagining responsible (technological) futures Organizers: Steffen Steinert, Martin Sand, Andrea Gammon, Filippo Santoni de Sio, Jordi Viader Guerrero, Aarón Moreno Inglés, Aafke Fraaije Background Engineers reshape the world through their design and operational decisions. Engineering ethics education, crucial in an era of rapid technological transformation, addresses the increasing societal demands for responsible innovation. As artificial intelligence and environmental challenges significantly impact our world, the responsibility of engineers extends beyond technology creation to ensure its benefit to public good and sustainability. In our attempts to respond to these demands in our technical university (TU Delft), we observe three major challenges that engineering ethics education faces: Challenge 1 -From problem solving to sense-making:Engineering ethics education, even when done in connection to concrete cases and problems, tends to focus on: a) problem-solving approaches, rather than broader problem-making and sense-making content; and b) methods based on abstract reasoning as opposed to embodied experience of the students. Current approaches are insufficient to develop moral and political sensitivity and motivation to act. Challenge 2 - Neglect of technological futures: The impact of technological futures on technological development is increasing but is currently neglected in engineering (ethics) curricula. This gives rise to the need to train engineers to responsibly assess and imagine their own technological and non-technological futures. Challenge 3 - More minds, fewer guides: With rising student numbers, there is a growing demand for ethics education at technical universities. At the same time, only limited personnel are trained in ethics and philosophy. Aim Our aim with this interactive experimental session is to introduce and discuss novel approaches currently being developed and implemented at Delft University of Technology to address these three challenges. Approach We briefly introduce the challenges identified, but use most of the session to focus on our approaches to tackling them. Participants choose between joining Activity 1, which responds to challenge 1, and engages them in actively experiencing power or social relationships mediated by (technical) systems, or Activity 2, which responds to challenge 2, where participants analyze selected existing futures according to a number of questions to develop their own vision utilizing different media. Rather than only showcasing existing, tried and true activities, we hope to enlist participants to help us refine and improve our works-in-progress. The second half of the session is an interactive discussion about the strengths and weaknesses of these activities, and about developing "Teach the teachers" strategies to tackle challenge 3: involving teaching personnel in other technical faculties in ethics teaching. Activity and Duration (min.) Introduction & current major challenges of engineering ethics education 10 Explanation of two activities & participants choose which to participate in 10 Activity/Challenge 1: "Experiencing space, power and social relations in the classroom" or Activity/Challenge 2: "Imagining responsible (technological) futures"

Behnam Taebi, Diana Martin	**11G EXP. SESSION**	Workshop: Reviewing for early career scholars - a bridge towards scholarly expertise and fair practice	Reviewing is an important part of the academic profession. It is essential for evaluating and strengthening the research conducted and submitted for publication by one's peers. But it also helps reviewers gaining a better understanding of the standards of academic publishing and developing themselves as authors. Being a fair and rigorous reviewer means being a valuable member of a discipline and academic community. Nevertheless, the process of reviewing can be daunting when first approaching it. The panel aims to elucidate what are the expectations regarding reviewing, and although its focus is in the area of science and engineering ethics, as represented by its key journal, the perspectives and advice shared can be translated for other philosophical disciplines. This workshop brings together editors from the journal Science and Engineering Ethics, with the aim of guiding participants towards the process of reviewing and presenting best practices. This is an active session that will stimulate a lively discussion between editors and participants. The workshop will have the following structure, combing presentation by the organisers of the workshop (members of the editorial board of Science and Engineering Ethics) with group and plenary activities: Introduction to the journal Science and Engineering Ethics by the editorial team, providing a description of the journal and its the role in developing this disciplinary field, the type of submissions welcomed, the publishing process from submitting an article to a decision and the role of reviewers in this process; Group activity, where in smaller groups oversaw by a member of the editorial team, the audience discusses a sample review, to identify the weaknesses of the review and ways to improve it; Plenary discussion, bringing together the overall weaknesses identified, followed by a general discussion about what makes a good review. The outcome is a listing of recommendations for reviewing fairly and rigorously; Conclusion, marking lessons learned and with the participants having the possibility to volunteer as reviewers. The target audience are early career scholars (PhD researchers and postdoctoral researchers) and researchers seeking to develop as reviewers. The session has three intended key outcomes. First, the participants will leave with an overview of the publishing process for the journal Science and Engineering Ethics. Second, the participants will gain an enhanced understanding of the reviewing process, captured through specific guidelines provided by the panellists. Thirdly, the broader academic community of philosophers and ethicists of technology, science and engineering, would benefit from the development of the new generation of reviewers guided by values of fairness and academic rigor, to ensure the growth of the field. This is especially important given the current challenges posed by a rise in retractions and the use of generative AI in editorial processes.
Joost Mollen, Elisabeth Does	**13C EXP. SESSION**	Transdisciplinary and participatory research and innovation: Philosophical questions and practical challenges	Motivation Transdisciplinary and participatory research and innovation are crucial in understanding, assessing, and designing responsible futures. These research and innovation efforts are often organized in so-called 'real-world laboratories', 'living labs', or other citizen science formats. In those research formats, academic, public, and private parties collaborate on developing and implementing solutions for grand social challenges, especially by developing and testing technological designs or policy interventions in real-world settings. These novel research formats also raise various semantic, epistemic, and ethical questions. What exactly are real-world experiments and laboratories? What are their epistemic characteristics? What ethical challenges do these new research formats bring about, and how can those challenges be addressed? How can we derive guiding principles for good transdisciplinary or participatory research that address epistemic and ethical aspects? Discussing these questions is essential to better understand, design, and govern the transdisciplinary and participatory research formats that offer foundational information on technological, economic, and social transformation processes. Transdisciplinary and participatory research formats are a recent phenomenon, and philosophically informed research on such formats is at a very early stage. Hence, to support the establishment of this field of research and communicate its relevance to philosophers of science and technology as well as 'living-lab practitioners, suitable conference formats are needed that foster exchange between pioneering scholars in the field. With our workshop, we wish to connect scholars who work on or are interested in philosophically informed research on transdisciplinary and participatory research formats. Workshop Outline We wish to provide an open discussion and group work-oriented workshop format (appr. 2-3 hours). We plan to form small groups which focus on the discussion of specific topics in the fields of ethics and epistemology of transdisciplinary and participators research formats. Group members then summarize what their discussion has revealed, especially in terms of current state of research, open questions and emerging issues. Group work results will be presented and discussed in a concluding debate. Overall, engaging in such a format shall help us to get an overview of current and emerging research questions, identify opportunities for collaboration, and work towards better integration and institutionalization of a young and still fragmented field of philosophical research. Furthermore, we encourage workshop contributors to also present research related to the topic of the workshop in general paper sessions or poster presentations at the conference. This way, exchange on workshop related topics can be deepened during the conference and become accessible to a broader audience. Finally, we will document discussion results of the workshop in a workshop report. We aim to publish this report in a suitable outlet. List of additional contributors (on top of workshop organizers E. Does & J. Mollen): - Jonne Maas (TU Delft) - Orlando Vazquez Villegas (TU Eindhoven) ... and all who join the workshop during the conference

Hans Dieter Kasperidus, Inga-Maria Eichentopf	**13E EXP. SESSION**	Interactive Exploration of Possible Climate Actions with the En-ROADS Simulator	Looking at the news these days, global reports related to wildfires, drought and floodings have become a new stable and establish a “new normal” of a world under a massively changing climate. As the 6th Annual Report (AR) of the Intergovernmental Panel in Climate Change (IPCC) states: With the global measures undertaken now to reduce the human greenhouse gas emissions, the global warming limit of 1.5°C, set in the Paris Agreement, will be reached already by 2030. And with that, the probability of extreme weather, putting the lives of people at risk and economies under pressure, will increase. Acting to mitigate the effects of climate change is needed more than ever. However, to communicate the problem in academic education as well as to a broad non-scientific and scientific audience is a critical issue. This also applies to the challenge to find reasonable measure to mitigate the effects of global climate change. For this purpose, we want to introduce the En-ROADS simulator as a supporting tool for high school and academic education and beyond. En-ROADS is a cutting-edge modelling tool developed by Climate Interactive (www.climateinteractive.org), a leading nonprofit climate & energy think tank, in partnership with the MIT Sloan Sustainability Initiative. The simulator is driven by a complex model that is based on a synthesis of the best available science literature on climate solutions. The intuitive graphical user interface is easy-to-use and 19 sliders allow testing interactively a variety of key processes in energy systems, economics, environment, and public policy in the global system and its connections to climate and many related fields. Thus, the simulator provides important insights in critical factors such as delay times, progress ratios, price sensitivities, historical growth of energy sources, and energy efficiency potential to the users. In the workshop we want to invite participants from different disciplines acting as expert team to test solutions and build a scenario for exploring climate change issues proposing global actions in the fields of e.g. energy efficiency, carbon pricing, reducing deforestation and carbon dioxide removal. The overall goal of the workshop is to create a scenario that limits global warming to 1.5°C above pre-industrial levels according to the 2015 Paris climate agreement. Thus, participants can experience in a facilitated group learning event the short-term and long-term impacts and connections of proposed actions on global temperature and other environmental, social and economic factors in real-time. As a result, participants will have developed a climate scenario that make connections between things they care about and the possibilities available to help ensure a resilient future. We expect from our experience in similar events that users will gain new insights and perspectives from the use of the simulator and we hope to get from this workshop valuable feedback on the usefulness of this approach for teaching and debating in a public context.
Tom Børsen, Diana Martin, Gunter Bombaerts	**9G PANEL**	Roundtable discussion: The role of theoretical foundations of the philosophy of technology and engineering ethics in engineering education	Engineering ethics education has emerged in the last decades as a discipline in its own right, seeing philosophers and ethicists expanding their work to comprise a pedagogical focus. It becomes increasingly important to reflect on how to bring topics of philosophical and ethical significance to the awareness of engineering students, both in terms of the theoretical lens employed and through methods to convey such focus. This roundtable session aims to foster an active discussion connected to the forthcoming launch of the Routledge International Handbook of Engineering Ethics Education, expected for publication in the second half of 2024. The volume includes contributions on both theoretical and practical themes ranging from foundational aspects, interdisciplinarity, teaching, and assessment in engineering ethics education. It is authored by 115 established and emerging scholars based on 6 continents. The roundtable will be led by three of the editors of the handbook (Tom Børsen, University of Aalborg; Diana Martin, University College London, and Gunter Bombaerts, TU Eindhoven), who are technology philosophers and ethicists with experience in teaching and researching engineering ethics education. They will be joined by two chapter authors (Gaston Meskens, University of Gent and Lavinia Marin, TU Delft) and one commentator (Carl Mitcham, Colorado School of Mines) for a conversation on the different theoretical foundations of the philosophy of technology and engineering ethics, as well as the need to expand this theoretical lens to non-western cultural perspectives and normative frameworks. Among the chapters highlighted in this roundtable discussion, we note the chapter 'A Post-normal Environment-centered Approach to Engineering Ethics Education,' coauthored by Shannon Chance, Tom Børsen, and Gaston Meskens, which contends that ethics is intricately intertwined with socio-ecological sustainability. The chapter lists several normative frameworks from sustainability science and environmental ethics. This intertwining goes beyond mere issues of knowledge, extending into the realms of politics and whose knowledge and values are deemed crucial in contemporary contexts. The chapter calls for a paradigm shift in engineering ethics education by introducing the concept of post-normal engineering (PNE) that has reflexivity as a way forward. Reflection is the focus of the chapter ‘Reflective and Dialogical Approaches in Engineering Ethics Education,’ authored by Lavinia Marin, Yousef Jalali, Alexandra Morrison, and Cristina Voinea. Drawing on Dewey's insights, the authors outline four key criteria for reflection in engineering ethics education: deriving meaning, systematic and disciplined thinking, communal interaction, and attitudes valuing personal and intellectual development. The roundtable speakers will provide examples of how theoretical frameworks in the philosophy of technology and engineering ethics can be adopted in teaching. Given that any account of the disciplinary foundations of engineering ethics must address where it is going as well as where it is, the speakers together with the audience will also discuss key recommendations emerging from the contributions in the volume. The audience will be an active part of this conversation.

Mark Thomas Young, Andrea Gammon, Ryan Wittingslow	**12C PANEL**	Retrofitting: Maintenance and Philosophy of Technology	Panel Proposal: Retrofitting: Maintenance and Philosophy of Engineering Maintenance is one of the fastest growing topics in philosophy of technology. This nascent field emerged from a recognition that existing work in philosophy of technology has tended to focus disproportionately on the design and creation of new things, at the expense of the widespread and diverse practices we employ to keep things going over time. Recently a growing body of philosophy literature has begun these forms of practice and exploring the ethical, ontological and epistemological issues to which they give rise. This aims to open up a new ground within this emerging topic by examining a specific form of practice through which artifacts are sustained through time. Retrofitting captures the practice of modifying artifacts after they are built in order to adapt them to emerging interests and needs and is commonly performed on an extremely wide range of different kinds of artifacts, from cars and computers to boats, bridges and buildings. Yet despite being a ubiquitous form of engineering practice, retrofitting has so far remained unanalyzed by philosophers of technology. This panel aims to address this neglect by bringing together three presentations which examine the practice of retrofitting through different philosophical lenses. The first presentation by Mark Thomas Young (U Bergen) aims to show how attending to practices of retrofitting promises to reveal political dimensions of artifacts which have been missed due to the traditional focus within philosophy of technology and STS on processes of innovation. The second presentation by Andrea Gammon (TU Delft) examines the nature of retrofitting as an environmental practice that connects the past, present and future in ways that have yet to be explored in either philosophy of technology or environmental philosophy. The third presentation, by Ryan Wittingslow (U Groningen) focuses on the significance of maintenance within the metaphysics of artifacts and aims to reveal how practices of retrofitting hold the potential to advance out understanding of the notion of function in exciting new ways. Together, these presentations aim to make an important contribution to the emerging topic of maintenance within the philosophy of engineering by revealing a variety of philosophical issues surrounding one of its undeveloped avenues: the practice of retrofitting. 1. Beyond Winner's Bridge: Retrofitting the Politics of Artifacts Mark Thomas Young (University of Bergen) Since the publication of Winner's influential article in 1980, the idea that artifacts have politics has remained a dominant theme in STS and the philosophy of technology. Yet despite exploring the political nature of artifacts from a variety of different perspectives, little of this work has paid attention to the politics underlying changes artifacts undergo after they are produced or constructed. In this presentation my goal will be to demonstrate how the neglect of such processes has obscured important dimensions of the political nature of artifacts, by casting them as stable features artifacts come to possess through design, rather than as ongoing processes that continue throughout the lives of artifacts. My goal will be to show how turning our attention to maintenance, and in particular practices of retrofitting, grants us insights into the relationship between technology and time that promises a major upheaval in how we conceptualize the political nature of artifacts. The first section of my presentation will address one of the fundamental assumptions behind Winner's argument: that artefacts acquire their political character through the formative activity of a designer. By revealing how artefacts often continue changing after their construction through the practise of retrofitting, I'll attempt to show that these formative acts are rarely confined to phases of production and construction and instead continue to apply throughout the histories of artifacts. In order to demonstrate this claim, I'll review a variety of different artefact kinds to illustrate the extent to which they continue to change after production. After detailing reasons why these processes of change should be considered forms of maintenance, I'll turn my attention to the political processes underlying such changes. In contrast to accounts in the philosophy of technology which depict breakdown and failure as objective states of an artifact, I'll demonstrate how, whether or not artifacts such as bridges require alteration in the form of repair or retrofitting, is an inherently subjective question that depends on values. Yet whose values take precedence in shaping the evolution of artifacts depends on questions of power. In the final sections of this presentation, I'll illustrate how attending to these processes encourages us to understand the politics of
Peter Pelz, Viet-Anh Nguyen Duc, Alfred Nordmann, Aleksandra Kazakova, Dazhou Wang, Andreas Brenneis	**12G PANEL**	The Multiple Languages of Engineering I	Organizers: Aleksandra Kazakova, STS, University of Chinese Academy of Sciences, Beijing // Alfred Nordmann, Philosophie, TU Darmstadt Recent years have seen many considerations of "Technology and Language". Some of these focus on the way we speak and think about technology, including the articulation of socio-technical imaginaries. Others work with the idea that the artful composition of (people and) things in a technical system or device follows a different kind of grammar in mechanical, electrical, and software engineering, that these grammars convey a sense of right and wrong ways of putting things together. Once one embraces the notion that technical system are composed and that Konstruktionslehren (construction manuals) initiate the apprentice into ways of doing things, it is only a small step to say that at the University engineers learn not only to solve technical problems but to express themselves. Engineering ideas are expressed in sketches, formulas, blueprints, pitches, prototypes. And once one considers technical devices, systems, or works as artfully composed or as the expression of an idea, we enter into the technical work as we do a literary or art-work, making sense, trying to understand, relating ourselves in a dialogical manner – in other words, we start engaging in hermeneutics. As with language-games, we are in the world by way of technology-games (as defined by Coeckelbergh). The proposed two-part panel will highlight various aspects of this interest in technology, language, and hermeneutics, emphasizing the multiplicity of languages. It includes six papers and two commentaries: Session 1 (1hr 45min): Presentations by Peter Pelz (20min plus 5), Viet-Anh Nguyen Duc (20min plus 5), and Alfred Nordmann (20min plus 5), commentary by Alexandra Kazakova (10min) plus 20min general discussion Peter Pelz (Mechanical Engineering, TU Darmstadt)The game carries further than the plan What is the function of a technical infrastructure? It channels and processes a resource so that it can be used to fulfil needs. A power plant, for example, splits the input energy flow into two output flows, namely the usable exergy flow and the anergy flow, which is regarded as wasted energy. The latter results in costs for society (social cost stream), while the former must be paid for by the user (private cost stream). The exergy flow at the outlet of the power plant is distributed by an electric grid or a mechanical or hydraulic gear to the usage devices such as a pump, a hammer, a spinneret and many other usage devices. The above text introduces terms such as needs, ideas, resources, flow, function, quality, infrastructure and cycle. It paints a picture and uses analogies between energy and data. It connects the technical sphere with the social sphere. By naming costs in two aspects, it combines economic and ecological thinking. This makes it clear to the author and the reader how language is shaped by images, analogies and associations. Technical design needs language, such as the language of (technical) drawings and symbols or icons, the language of mathematics for physical models of components represented by symbolds, programming languages for interaction with machines, the language for semantic (knowledge) graphs based on subject-predicate-object based connections. The latter is used for communication from machine to machine, from machine to human and from human to human. At TU Darmstadt, for example, we are transferring game-based negotiations between people to negotiations between technical agents. Thus, a pump is not just a pump, but an agent that trades promises, functions, has qualities and budgets. The trading partner is the valve adjacent to the pump or a neighboring pump. Both negotiations take place on marketplaces with designed rules.This gives riseto questions regarding the interaction between humans and machines. If we look at the usability of languages between humans at one pole and, say, CPIs at the other pole, it is quite interesting how we can classify languages in terms of their ability to express ideas, functions and qualities. Finally, the advancement of language and technology requires both creativity and openness. Is evolution hindered by the fact that everything should be machine readable? Viet Anh Nguyen Duc (Philosophy, TU Darmstadt)Unveiling Technology's Essence: A Hermeneutic Journey through the Debate on Technology Hermeneutics In recent decades, numerous contributions from the fields of philosophy of technology and STS have appeared under the methodologically guiding term of technology hermeneutics. They have in common the core idea that

Andreas Brenneis, Dazhou Wang, Alexandra Kazakova, Alfred Nordmann, Peter Pelz, Viet-Anh Nguyen Duc	**13G PANEL**	The Multiple Languages of Engineering II	Organizers: Aleksandra Kazakova, STS, University of Chinese Academy of Sciences, Beijing // Alfred Nordmann, Philosophie, TU Darmstadt Recent years have seen many considerations of "Technology and Language". Some of these focus on the way we speak and think about technology, including the articulation of socio-technical imaginaries. Others work with the idea that the artful composition of (people and) things in a technical system or device follows a different kind of grammar in mechanical, electrical, and software engineering, that these grammars convey a sense of right and wrong ways of putting things together. Once one embraces the notion that technical system are composed and that Konstruktionslehren (construction manuals) initiate the apprentice into ways of doing things, it is only a small step to say that at the University engineers learn not only to solve technical problems but to express themselves. Engineering ideas are expressed in sketches, formulas, blueprints, pitches, prototypes. And once one considers technical devices, systems, or works as artfully composed or as the expression of an idea, we enter into the technical work as we do a literary or art-work, making sense, trying to understand, relating ourselves in a dialogical manner – in other words, we start engaging in hermeneutics. As with language-games, we are in the world by way of technology-games (as defined by Coeckelbergh). The proposed two-part panel will highlight various aspects of this interest in technology, language, and hermeneutics, emphasizing the multiplicity of languages. It includes six papers and two commentaries: Session 2 (1hr 45min): Presentations by Andreas Brenneis (20min plus 5), Dazhou Wang (20min plus 5), and Alexandra Kazakova (20min plus 5), commentary by Alfred Nordmann (10min) plus 20min general discussion Andreas Brenneis, Arthur Wei-Kang Liu, Jörn Wiengarn (Philosophy, TU Darmstadt)What were the Club of Rome Reports meant for? Some ideas concerning their hermeneutical analysis In the seminal reports to the Club of Rome in 1972 ("Limits to Growth") and 2022 ("Earth for All"), rhetorical strategies were employed alongside ostensibly objective computer models to present and argue for Responsible Futures. Rather than assuming that both reports pursue the same purpose, our presentation adopts a hermeneutic approach to interpret the underlying intentions and aims of each document. The hypothesis we aim to elucidate is that due to the world's development as well as the discursive uptake of the first report, substantial changes in the communicative aims were implemented. Against the backdrop of both reports relying heavily on models, we question the relation between the objectivity that is claimed by applying models and the rhetorical elements that form substantial parts of the reports. Key to our investigation are the rhetorical registers employed in the reports. We aim to scrutinize the use of metaphorical language, the tone of address to readers, and the incorporation of technical jargon, the range of scenarios depicted (positive, neutral, negative), and how they are portrayed differently. A main focus of our investigation will be the construction of agency within the reports, identifying the active and passive agents of change and interrogating the roles ascribed to scientists and scientific knowledge. By delving into the portrayal of urgency and the attribution of responsibility, we seek to elucidate the underlying motivations and purposes driving these influential documents. Through our analysis, we aim to uncover the pragmatic functions of the reports' future-oriented discourse: from informing and emotively engaging readers to advising on potential courses of action. Ultimately, our hermeneutic analysis offers a nuanced understanding of the Club of Rome reports, illuminating the complex interplay between rhetoric, scientific modeling, and the construction of visions for the future.
Mark Bessoudo	**EXP.1**	Google Street View: Ethics, Creative Freedom and the Future of Street Photography	"Today everything exists to end in a photograph." — Susan Sontag, New York Review of Books, 1974 Google Street View is a popular online mapping tool integrated with Google Maps that provides users with a 360-degree street-level view of city streets, back alleys and dirt roads around many parts of the globe. While it provides users with unprecedented utility, it also raises some unique philosophical questions about ethics, creative freedom and the future of technology. Google compiles their Street View image database by using their own proprietary vehicles mounted with special cameras which collect images every few meters. Billions of photos are then stitched together to create a vast visual library capturing the elements of the built environment: people, buildings, signage and infrastructure. Like more traditional street photography, Street View serves as a time capsule into public life by capturing the banal and mundane moments of everyday ordinary life in the process: a roadside market in Daka, Senegal; the densely packed Shibuya district of central Tokyo; humble mountain villages of Bhutan; tiny chapels in fishing outposts dotted along the Greenland coast. It is for this reason that Google Street View has amassed a legion of fans. But it also raises some unique philosophical questions about ethics and creative freedom. If using Google Street View as a surrogate for "street photography", are the screenshots captured from it a form of aesthetic consumerism? After all, "to photograph people is to violate them," claimed the cultural critic Susan Sontag in her 1977 classic On Photography. "It turns people into objects that can be symbolically possessed." Are millions (billions?) of people "symbolically possessed" by Google's Street View database? Furthermore, has creative freedom simply been outsourced to an algorithm? Art is always a balance between the tensions of freedom and constraint. With Street View photography users have ultimate freedom in that they can visit almost any street on Earth, but are limited to capturing only what Google's cameras have – for example, you cannot achieve better lighting, get a different angle or wait for a more interesting subject to wander into frame. You must work with what you've been given. This experimental art exhibit will explore the nature of these ethical and creative questions. Through a series of Google Street View "photographs" that I have compiled from around the world, this exhibit will invite viewers to consider the ethical implications not just of traditional street photography and Google Street View of the past and present, but of the coming (and inevitable?) introduction of far more advanced and intrusive technologies such as Artificial Intelligence, Augmented Reality, drone photography and 24/7 mass surveillance. The exhibit would be accompanied with introductory text explaining my motivations for this project. While Sontag's observations were meant as a critique of the growing aesthetic consumerism of the 1970s her criticism still resonates, perhaps even more so today: "The camera makes everyone a tourist in other people's reality, and eventually in one's own."

Carolina Ibarra Castro **EXP.2**	Art exhibition about ethics, autoperception and representation of indigenous peoples in image generation programs with artificial intelligence	[See corresponding mail and PDF sent by author on april 16] This proposal for fpet 2024 at ZKM is an artistic exhibition that seeks to make visible the biases and prejudices of representation and the digital self-perception of indigenous peoples within the digital landscape generated by an artificial intelligence and machine learning language model, programs where the user introduces a prompt or description in natural language, and from where an image emerges from an artificial intelligence model. With a focus between engineering philosophy - centered on the algorithm - and STS studies, this work addresses the digital self-perception by users of indigenous peoples in machine learning programs and, in turn, the cultural biases and prejudices that images from these programs, such as Gencraft, Dall-E and other open access programs, promote reflection from the hypothesis that for the global south, colonialism continues in the digital world, while it is from the north that are generated the data models and algorithms that shape an image of Latin America, just like the drawings of explorers from the early 19th century. The exhibition exhibits visual material from the interaction with indigenous communities in Chile and images generated by artificial intelligence programs. It is composed of at least 10 small screens with testimonies and data visualizations on the one hand, in accordance with the ZKM exhibitions, and printed images as photographs, with the proposal - not yet realized - of objects that function as visualizations of 3D data. This exhibition is part of a digital ethnography study to investigate how generative art platforms interpret and generate visual content when provided with some keywords, to explore how concepts about the notion of Latin American indigenous communities are interpreted, and is approved to be published as book chapter in the Practicing Digital Ethnography issue. (2026, Routledge and Elon University, North Carolina. Devin Proctor Editor, Assistant Professor of Anthropology) Usually, and as has been discussed with local communities, indigenous representation in programs like Gencraft is cartoonish. To promote the financing of this exhibition, we request a letter of acceptance from the forum, thank you. By unpacking the biases present in generative art, we aim to foster a more inclusive, culturally sensitive, and equitable approach to AI-driven creative expression. Bibliography: Zou, J., & Schiebinger, L. (2018). AI can be sexist and racist—it's time to make it fair.

3) Posters

Lee-Ryeok Han	A Two-timed Stratgy for State Identification	A two-timed strategy is studied for a being and interactions with other individual beings. A teleological philosophy from the Aristotelian paradigm insists that the future decides the past which is thought to be the source of potentiality. A teleological philosophy allows to move backwards in time flows. As a result, the Aristotelian paradigm seems to have something to do with quantum phenomena. The Aristotelian paradigm requires movements to state the a posteriori and the a priori beings. However, the Aristotelian paradigm argues that a subject instantaneously exists temporal so that no time flows. In a nutshell, the Aristotelian philosophy sheds light on its paradigm in time flows with a spontaneous temporal subject in no time flows. Regarding quantum phenomena, this contradiction seems to inevitably leave room for questions on the a priori being to single out among the a priori ontological candidates. Aristotle argues that being is supposed to consist of spatio and temporal identities which inevitably on the move, nonetheless, a subject is temporal. Thus, they are considered to provide the basis for respective a spatio object to be observed and a temporal subject to observe for this research. Self-identification could be realized when a being observes itself as an object wherein a subject is not an outer observer. Self-identification is explained by using the similar concept of Out-of-Boddy experiences in no time flows so that a subject is temporal. Hence, self-identification in no time flows is assumed to depict the Aristotelian paradigm on Feynman diagram which is not realizable in time flows. When self-identification is realized in time flows, versatile ontological beings on the same time stamp would be elucidated such as entangled or superposed state. Therefore, two-timed is introduced to realize self-identification in time flows on Feynman diagram. Specifically, the notion of dream is adopted for two-timed flows. Accordingly, the Aristotelian paradigm is modified to include the concept of dreaming between the future and the past to realize self-identification in time flows and defend the teleological determination. Unfortunately, two-timed on Feynman diagram does not distinguish anything about the versatile (entangled or superposed) a priori being but specifies it with respect to self-identification in time flows. As a result, a two-timed strategy expands studies on a being to interaction with many other beings who have their own individual subjects and objects. In other words, there are many beings who simultaneously observe and are observed each other. As Moore claims, this epistemological expansion can be explained as 'How do we know that there exist any other people who have perceptions in some respects similar to our own?' The interaction through two-timed flows not only converts being who is bosonic to fermionic or vice versa, and then decides what is being observed. Finally, a two-timed strategy is adopted in the Kalman filter which is one of the best stochastic algorithms in the field of sciences and engineering for engineering realization. Han, L R 2021 Two-timed: Self-identification Strategy. Social Phenomena – International Research Journal, Volume 1, No. 1: 75-78. DOI: https://doi.org/10.47929/2305-7327_2021.01_75-78

Johanna Teresa Wallenborn, Maximilian Roßmann	The Metaphorical Landscape of Sustainable AI: Exploring technology Futures on Twitter	This multimodal metaphor analysis of the "Sustainable AI" discourse on Twitter uncovers the absence of a unified vision. Instead, the term functions as an umbrella term, attracting diverse actors and shedding light on various aspects of sustainability and AI. Using a dataset of 2,367 tweets collected via Twitter API 2.0 from 2017 to May 2023, the study analyzes tweets that received more than 15 likes (n=84), following the methodology of systematic metaphor analysis by Schmitt et al. (2018). Our analysis and discussion of the different metaphorical interpretations of both, AI and sustainability, indicate that "Sustainable AI" underscores the transformative potential of integrating both concepts, yet also risks oversimplifying complex environmental issues, covering black-box applications, and promoting techno-solutionism. The study concludes that while "Sustainable AI" can indeed harness the ongoing AI hype to prioritize sustainability, a critical examination is essential to navigate potential pitfalls when confronting the climate crisis.
Michael Mayer, Bettina Kamm, Jan Rabold	Approaches to increasing explicit opportunities for the acquisition of professional ethical competencies in higher education	By acquiring professional ethical competencies during their studies, universities ensure that engineering graduates (i) act responsibly, (ii) safeguard the integrity of the profession, and (iii) contribute positively to society. Despite the importance of professional ethical competencies, explicit ethical discourse only takes place to a limited extent during students' higher education. Therefore, students often acquire ethical competencies unaccompanied, for example by interpreting and adapting the behavior of lecturers or researchers. By further developing explicit ethics-related learning opportunities, students can be even better prepared for the responsible handling of tasks and challenges in society, economics, and science. This requires learning settings, which support (i) an understanding of the impact of technological developments on the environment, society, and the economy and (ii) at the same time the ability to integrate ethical principles such as a sense of responsibility, fairness, and consideration for future generations into professional work. The further development of explicit ethics-related learning opportunities can better ensure that engineers are not only professionally competent, but also ethically reflective individuals, who also contribute positively to democratic education and sustainable development. In addition, such qualified students can also make a better contribution to the further development of the higher education teaching and learning system during their stay at the university. The discourse on currently highly relevant future skills (e.g., assumption of responsibility) and the need for further development of curricula as a result of changes in the general higher education framework (e.g. AI, education for sustainable development) allow for a synergetic increase in ethics-related teaching and learning. In order to receive feedback for the approaches currently being discussed in the teaching unit "Geodesy and Geoinformatics" (GaG) at the Karlsruhe Institute of Technology (KIT) and to increase richness of perspectives in the change process, selected learning settings of BSc degree programs will be presented at fPET2024. In particular, the two well-established approaches "Project Introductory Phase" and "Team Teaching" will be discussed. The thematically focused "Project Introductory Phase" (e.g. 2022: Professional Ethics, 2023: Education for Sustainable Development) takes place during the second semester and enables students to (i) engage in a topic-related exchange with professionals and (ii) explicitly reflect on ethics-related future skills relevant to their professional field. During "Team Teaching", students in the sixth semester are mentored by GaG-lecturers and ethics experts in a seminar-style teaching format dealing with ethics-related topics (e.g., AI, tracking of persons).In addition to these two well-established learning formats, during fPET2024 a recently developed lecture (title: Empowerment for the Future and Society), which will prepare students even better for their important future role in rapidly changing systems (e.g. professional field, society), will be put up for discussion.
Diana Martin	Integrative engineering education: exploring the responsibility of technological universities	The grand societal changes and challenges of our time demand a broader role of technological universities, thus opening the question of how this role has evolved over time and how to understand their current responsibility. We argue that not only technological artefacts are not neutral, but also technological universities are not neutral: through their educational activities, institutional practices and programmatic documents, mottos or vision statements, technological universities embed specific values and convey specific understanding of what engineering is. Following a literature reveiw of the history of engineering education, the poster presents 5 major orientations of technological universities and their associated responsibilities, that can better help understand the ideal of an integrative engineering education. As such, an integrative model of engineering education can be understood as comprising scientific, market, policy, societal and ecological components. Each of these are linked with a specific definition of responsibility. The paper contributes to recent debates about the role of engineering education, by putting forward an instrument which can serve as a "transdisciplinary passport" that technological universities can use to map their practice and articulate their identity and associated responsibilities.
Nick Treanor	The Epistemic Role of Testing and Certification in the Construction Industry	This poster would be a joint presentation by an academic philosopher working at a major UK research university and a senior engineer in the UK who is a former vice president of the Institution of Civil Engineers and has lead teams engaged in multibillion-pound engineering projects. It would focus on the epistemology of engineering, specifically with regard to the epistemic role of testing and certification within the industry. Many core notions that underpin product certification regimes are ones that philosophy specialises in; for example, the relation of evidence to theory and the role of expertise and testimony in knowledge transmission. But also, and more generally, engineering is a precision discipline and philosophy is precise in ways that engineering is vague. In this paper, we will focus on developing and bringing clarity to three ideas that bridge philosophy and engineering: 1. The epistemic and practical consequences of the tendency of regulatory bodies, testing regimes, and industry practitioners to treat qualities, characteristics and phenomena that are graded or continuous in quality as if they were threshold qualities. This leads to a 'just good enough' mindset/practice, which gives rise to several problems: (a) it's worse if things are just good enough, vs ever closer to excellent, (ii) aiming for just good enough will inevitably yield a distribution of outcomes clustered around just good enough, some a bit above and some a bit below – that is, not good enough, and (iii) even if one aims and hits just good enough, there is no robustness, so small problems or changes to context can bring things below the threshold, which gets discovered only by the occurrence of a failure. 2. The epistemic and practical consequences of conflating idealisation and reality, particular in testing and certification; this section will discuss how this played a central role in the 2017 Grenfell fire in London, drawing on material presented at the Grenfell Tower Inquiry. Test results tell you how a product performed in a testing environment, not how they will perform in the real world environment, which will be similar to, but also different from, the testing environment in innumerable ways. But the industry tends to treat test results as testimony to the effect the product will be safe in the built environment. That is, they treat test results as showing or demonstrating real world safety, rather than as being an evidence source that, when combined with other knowledge, allows someone to make an accurate judgement about the performance in a given real world situation. 3. The epistemic and practical consequences of a reliance on testing displacing, rather than aiding, expert judgement, due to a compliance mindset and a misunderstanding of what testing does and does not show. Rather than increasing understanding, current regulatory regimes tend to undermine it in a way that undercuts rational endorsement of standards and motivation to fulfil them. We will discuss this with particular reference to code-stretching in construction.

Florian Richter	From Human-System Interaction to Human-System Co-Action: Ethical Assessment of Generative AI	Technology ethics has become a multifaceted research area ranging from normative to empirical investigations. Specifically, in the field of human-machine ethics, it is paramount to investigate technology's mediating role and its ethical implications. The conceptual level has been investigated, for instance, in postphenomenology. [1], [2], [3] However, in recent years, several studies have been conducted to shed light on the influence of technology on human behavior in human-machine ethics, from the influence of chatbots [4], to the ethics of autonomous vehicles [5], [6], hybrid traffic [7], and the ethical implications of care [8]. However, immersion in virtual reality and real-time feedback from systems to subjects via Generative AI is still in the making and is discussed under the concept of "co-action." The ethics of human-system interaction, which has already been disrupted by AI technologies, faces new challenges in developing ethical assessment tools for human-system co-action. Aligning expectations of systems such as conversational agents (CA) and humans has been discussed under a "mutual theory of mind." [9], [10] It should allow "smoother human-CA conversations" by adapting the mutual theories or models of the mind. [9] However, it is unclear if such models of the mind reflect either the system's functioning or the human agent well. It might be doubted whether something like this can be interpreted as a (mutual) theory of mind where the system develops a model for the user, and the user has a model of the system. Instead, the system ascribes roles, creates relevance rankings, establishes routines for the user, and based on such an analysis, the system offers recommendations, explanations, or options for the user. By having an actual theory of mind, the user can also, to a certain extent, exploit the system's strategic functioning to reach a specific goal or objective. This could be extended by assuming that the system or the virtual conversational agent has adaptational strategies, too, so that it can react to these strategies. Hence, game theory should be considered to further investigate how humans and systems employ strategies to get what they want, need, and seek depending on their motivations or computational structures. The poster will show four contributions. First, the problem situation of human-system-co-action is outlined to assess and summarize the ethical implications of the role of technology. Second, the methodological approaches of a "mutual theory of mind" and "parallel communication" are contrasted, and the shortcomings of the former are summarized. Third, the creation of user modeling is analyzed to demonstrate the strategies of systems. Furthermore, some examples of generated pictures are integrated. However, and that is the final aspect of the poster, such strategies need to be discussed in relation to conceptions of fairness. Moreover, a package of measurements is outlined to show further steps of empirical research in this context. References [1] P.-P. Verbeek, "Beyond interaction: a short introduction to mediation theory," Interactions, vol. 22, no. 3, p. 26–31, May - June 2015. [2] R. Rosenberger and P.-P. Verbeek, "A Postphenomenological Field Guide," in Postphenomenological Investigations: Essays on Human-Technology Relations, R. Rosenberger and P. Verbeek, Eds., New York/London, 2015, pp. 9-41. [3] I. van de Poel and P.-P. Verbeek, "Can technology embody values?," in The Moral Status of Technical Artefacts. Philosophy of Engineering and Technology, vol. 17, P. Kroes and P. Verbeek, Eds., Heidelberg/New York/London, 2014, pp. 103-124. [4]
Elisabeth Does	Ethics in Transdisciplinary Research Formats: Reflection Guidelines for Responsible Research Practice in Real-World Labs	The project I would like to present at fPET is dedicated to ethical challenges in so called real-world labs, also known as living labs. Real-world labs are transdisciplinary and participatory research formats in which academic, public, and private parties collaborate on developing and implementing solutions for grand social challenges, especially by developing and testing technology designs or policy interventions in real-world settings. Due to their methodological peculiarities at the intersection of science and society, real-world labs come with ethical challenges that differ from those in more traditional research formats. Hence, the need was identified to develop ethics guidelines which address the specifics of transdisciplinary and participatory research in real-world labs. In 2022, a team of real-world lab practitioners and philosophers at the Karlsruhe Institute of Technology (KIT) began working on a code of ethics for real-world labs with the intention to particularly serve community needs of the "Netzwerk Reallabore der Nachhaltigkeit" (Network of Real-World Labs for Sustainability) in German-speaking countries. In the course of this project, the need for ethics guidelines which serve a broader audience – including real-world labs which do not conform with specific normative orientations pre-defined by the aforementioned network – was identified. Furthermore, pursuing variety in guideline formats was considered helpful in order to overcome limitations of a classical code of ethics. Promoting a fixed set of values for a specific community, such a code would especially serve as a basis for identity development in this community and not so much as a tool for self-empowerment in reflection processes. In effect, the project at KIT was divided into two sub-projects. One dedicated to the originally intended "Code of Ethics for Real-World Labs for Sustainability", and a second one dedicated to differently designed "Reflection Guidelines for Responsible Research Practice in Real-World Labs". At fPET, I would like to present those "Reflection Guidelines" in a poster session. I would like to explain how they represent an example of research ethics understood as an on-the-job reflection practice available to all researchers, rather than a catalogue of fully explicated values, pre-fixed by those who establish them for those who shall adhere to them. I would like to shortly present the development process of those guidelines and explain some of the most distinctive characteristics of the format. For example, I would like to explain in more detail the ethically relevant fields of actions in real-world lab research along which the guidelines are structured. Special attention would be given to their core feature, namely comprehensive lists of questions for self-reflection which are given for each field of action. Those questions are designed such that they may empower real-world lab practitioners to independently and systematically reflect on ethical challenges specific to their own real-world lab project. With my poster, I would like to contribute to debates concerning ethical challenges in transdisciplinary research and technology development as well as to debates concerning the development of different formats of ethics guidelines.
Anja Bodenschatz	Does the Implementation by Autonomous Systems Make Randomization in Ethical Dilemmas More Acceptable?	Autonomous systems based on artificial intelligence (AI) are becoming integral to society. This raises questions concerning ethical dilemmas that have remained unresolved in the societal debate. Ethical dilemmas of existential threat, such as life-threatening resource shortages, have become rare in our society and are left to individual actors or institutions to decide on a case-by-case basis. With the necessity to program autonomous systems, this is no longer possible; scenarios must be anticipated, and guidelines for dealing with them must be programmed in advance. Concerns arise about whether machines can and should make these ethical decisions. AI technology uses a mathematical function to optimize actions towards a target objective. AI will, therefore, always statistically discriminate in dilemmas, depending on the objective. In many democracies, however, this is ruled out or at least contested by anti-discrimination laws. The studies presented here explore the possibility of addressing ethical dilemmas through random draws, in which possible costs and benefits are randomized between affected people. This practice has been used and has found societal acceptance repeatedly throughout history but is largely excluded from the current societal debate on dealing with ethical dilemmas. Today, randomization is a taboo topic in commonsense morality where ethical dilemmas of high impact are concerned. However, with the need to program autonomous systems free of discrimination and with modern machines' technical possibilities and limitations, it is gaining importance in the philosophical and behavioral ethics debate. Several empirical studies found that randomization in ethical dilemmas may be accepted by many if applied by autonomous systems. The empirical studies presented here explore what may drive a higher acceptability of randomization in ethical dilemmas when implemented by autonomous systems rather than human actors.

Schultz M.A., Maximilian	And now we're doing such an AI project: On the State of Non-Governmental Organizations' Engagement with Artificial Intelligence - Opening a New Field of Research	Non-governmental organizations (NGOs) play a central role in civil society (Hall & O'Dwyer 2017, Pogorelskaja 2015) by interacting with governments, businesses, and other societal actors. Thus, NGOs are an important actor in civil society, representing many groups with a focus on the common good. Artificial Intelligence (AI) is a currently omnipresent topic. While technology corporations and businesses are driving the development of AI, the question arises to what extent civil society, and specifically NGOs, could benefit from the new possibilities that AI brings. The methodological approach to answer this question relies on a Mixed-Methods approach (Kelle 2014), in which data is collected and synthesized in four phases. In the first phase, a scoping review (n=14) was conducted, and in the second phase, an initial exploratory survey with NGO experts (n=5) was carried out. The third phase involves a large quantitative survey. This is followed by in-depth interviews. The first two phases of the project are already completed; the results can be presented in the form of a poster. The scoping review aims to analyze (non-)scientific literature on NGOs and AI; it was conducted using a search string with German and English terms in the databases Web of Science, WISO, and Science Gate. Both scientific and non-scientific texts were included. The initial results show that a large portion of the identified sources consists of press releases. Only two scientific studies and one research report that includes cross-references to the topic were found. The texts deal, on one hand, with theoretical considerations of AI in the context of NGO work, for example, the necessary legal regulations. On the other hand, initial applications, mostly in the health sector, are presented. The texts often deal more with theoretical considerations of AI in NGOs and less with applications. In five exploratory interviews with experts or employees of NGOs, knowledge, acceptance, needs, and risk assessments were collected. Attention was paid to different sizes and fields of action on the part of the NGOs. The interviews show that AI is currently not a priority for NGOs, but there is interest in the possibilities. While the interviewed NGOs have carried out initial projects and trials and rated the experiences as positive, there are still no established structures or clear ideas for the use of AI. All interviewed organizations have already made initial experiences, especially with Large Language Models like ChatGPT. This technology is seen as a potential solution for structural challenges, but its use is limited to smaller applications or funded AI pilot projects. The NGOs rate the current use of AI as positive. The findings of the survey will flow into the questionnaire design of an online survey among NGOs in the third phase. In the fourth phase, in-depth interviews with NGOs will be conducted. In conclusion, the limited number of identified texts and the statements of the respondents indicate that the topic of AI in NGOs represents a young field of research and application. The results of the scoping review and qualitative interviews can provide an important basis for further research. It becomes clear that NGOs are mostly likely just beginning in terms of using AI. Although the interviewed NGOs evaluate the use of AI for their own purposes as positive, the responses of the interviewees clearly indicate that there are hurdles and fears, for example, in the application of AI in the social or health sector, which must not be overlooked.
Heesbeen PhD Candidate, Johannes Willem	Design Practice as a test-bed for Institutional Logics	In the institutional logics perspective, Design Practice represents an exaggerated cosmology of phenomena. Since design and design practices are product of and instrumental to cultural reproduction, the institutional tangle that is manifested seems quite unique. Although design practice intersects many institutional logics and patterns, the relative autonomy of intersectional fields remains vital (Ocasio 2023). In the institutional perspective, more and less articulated and entangled systems are not merely aggregations of individual or social behaviour and practices but remain emergent structures that maintain causal effects over lower-level structures. Situated in the lower-level structures, including groups and individuals, causal effects work back into higher levels, in fact, the lower-level clusters are critical in the continuation and reproduction back to the higher-levels. Stability and permanence at higher levels, societies and the world system, tend to be more general and abstract in comparison to more specific behaviours and practices observed at lower levels of analysis. In effect, institutions are sociocultural constructions that provide models for organizing, and therefore are always present but hardly ever in the foreground of organizational phenomena. Ocasio (ibid) emphasizes how micro and macro-institutions are not disjointed polarities, but reciprocally shaped by the sociocultural phenomenology of networks of committed practitioners, embedded across levels of analysis - organizations, organizational fields, communities, societies and the world system. Embeddedness in each and the other. Seeing institutional logics and embeddedness in the reorientation on design proposed by Johan Redström, “a design theory that is inherently unstable, fluid and dynamic in character”(Redström 2017, 2) is the central premise to be studied. Furthermore, considering design in the category of communicative activity, seeing communicative actions as the basic element in the social construction of what constitutes our interests and passions, design is instrumental in materialisation and objectification of any kind, both mediated and direct. To work in the dynamic of practice, the view ‘from the inside’ can be found in the life-world perspective of Schutz, “knowledge has no given structure and is in continual flux, so that at any given moment its characteristics are reformulated as they are called upon by the demands of any particular NOW” (Schutz 1953).
Oliver Shuey	What can post-phenomenology tell us about engineering knowledge?	Engineering practice in the US is deeply tied to philosophical conceptions of technology. Taking the engineer to be an arbiter of technological change, the engineer is partially responsible for shaping what technology is and ought to be. Engineers are also humans with limited perspectives, their standpoints are subjective and may not be suited for solving every problem they encounter. At the same time, engineers make use of operational design principles that force designers to essentialize problems. They may even fail to implement more advantageous solutions based on inadequate knowledge of or experience with users of their technical artifact, as is the case with pieces of adaptive tech for disabled people. As a result, engineers are limited in what they can know and how they can apply their knowledge. However, engineers still foster intimate relationships with the technologies they use and make because of their repeated interactions with artifacts. The engineer gains design experience throughout their career, contributing to the knowledge base they may pull from for future technological development. Engineering, in a sense, is an embodied experience; they learn how to do engineering through interactions with technology. Post-phenomenology may serve as a useful analytical tool for studying engineering and the development of engineering knowledge. By utilizing deeper insights into and descriptions of the human experience, post-phenomological analyses explore the ways in which technology shapes our choices and worldviews. As socio-technical problems become increasingly “wicked” and global, building more just artifacts requires that epistemic gaps in engineering knowledge are identified and addressed. It may be possible to gain insights into the development and maintenance of engineering knowledge via post-phenomenological accounts of interactions between engineers and their artifact that are not evident through less intimate conceptions of human-technology relations. This poster explores what a post-phenomelogical investigation of engineering may entail, the limitations of such analysis, and how this information can be used to address shortcomings in engineering.

Katja Nau, Christoph Steinbach, Harald F. Krug, Dana Kuehnel, Alexis Bazzanella, Matthias Finkbeiner, Jessica S. Hoffmann, Andreas Mattern	MANTRA - data on innovative materials for sustainability and transfer	The development of new technologies, e.g. for energy saving and climate protection, represents some of the key issues of today. Research focuses on new materials and catalytic processes that contribute to saving energy or protecting the environment, both in technical applications and in consumer-related products. The Federal Ministry of Education and Research (BMBF) is funding scientific projects developing such promising materials and processes. With the implementation of the "From Material to Innovation" framework programme, the BMBF is creating a funding opportunity for the development of innovative catalyst and membrane materials whose novel or improved use is intended to make energy- and CO2-rich processes more efficient and sustainable with the "Materials for Process Efficiency" funding guideline. Scientific support for funding measures by projects that take on subject-specific networking, cumulative science communication and science-related coordination of activities creates added value compared to the funding of individual, non-networked and unsupported scientific research, and development projects. Accordingly, the project MANTRA - data on innovative materials for sustainability and transfer – funded by the BMBF, is a scientific communication project for networking, public relations, and the development of indicators for sustainability as well as industry/practice transfer. MANTRA supports the material hub initiative "MaterialNeutral - Materials for Process Efficiency", presenting knowledge on health effects and sustainability on novel materials and processes related to catalysis and membranes to the public. The aim of the material hub initiative MaterialNeutral is to develop mission-orientated material innovations for pressing social and industry-relevant issues using a holistic and cross-actor approach. Cooperation and synergy potentials between science, industry and society are to be utilized in terms of sustainability and resource efficiency. MANTRA will communicate these topics via a website (www.materialneutral.info), brochures and information flyers using an interdisciplinary approach. Methodological expertise in Germany is to be further expanded as part of the research and development work. The utilization and scaling of the results up to technology transfer into industrial applications are an essential part of the hub concept and should already be established as part of the funding projects and driven forward over the funding period. Cooperation and synergy potentials between science, industry and society are to be utilized for the desired sustainability and resource efficiency. In this way, innovative materials and substances are to be made safe, available, environmentally friendly, and sustainable. In this context, MANTRA will facilitate the knowledge transfer on materials safety and sustainability among all relevant actors from material developers to consumers. In addition, novel methods and indicators will be developed for the assessment of novel materials sustainability. Supporting the generation of networks between the research projects is envisioned to facilitate the development of novel and environmentally friendly materials in Germany. The project is funded by the Federal Ministry of Education and Research (MANTRA Data on innovative Materials for Sustainability and Transfer, 03XP0583).
Andreas Lösch, Janine Gondolf, Christian Büscher, Ulrich Ufer	Transformation Assessment – observing and (co)shaping sociotechnical transformations	Transformation programs are being launched around the world to address global challenges and enable a more sustainable future. However, each transformation requires a reorganization of basic socio-technical systems that can affect all social subsystems. These changes have far-reaching consequences and affect actors who may not be directly involved in the transformation effort. Crisis-driven transformation programs, such as those focused on energy or digital transitions, can cause significant disruptions in social coexistence because they are often initiated without consideration of the necessary social and structural changes. These linkages and interdependencies within transformation activities have received limited attention, despite their importance. Understanding them is critical to the success and effectiveness of transformation. Technology Assessment (TA) is involved in transformation projects at all levels (Grunwald 2019). TA's expertise lies in its interdisciplinary and integrated range of theories, methods, and processes to critically anticipate, examine, evaluate, communicate, and help shape (emerging) transformations, similar to other approaches (e.g. in STS, Konrad et al. 2016). Rather than primarily addressing scientific questions, TA is concerned with creating structures that generate solutions to problems. It places these activities in a larger societal perspective, assessing their potential for change and their impact on social sub-systems, while at the same time considering their visionary promise and their interconnectedness. With our poster we want to outline a concept for "transformation assessment" by and in TA. It is intended to help guide transformational projects and to make their possibilities and consequences visible, assessable and reflectable. It integrates (a) the analysis of transformations from a systems-theoretical perspective (Büscher 2018), (b) an immersive anthropological and cultural studies perspective (Ufer/Hausstein 2021), and c) the perspective of vision assessment (Lösch et al. 2023; Schneider et al. 2023). In this sense, "transformation assessment" could be a resource for both non-TA and TA transformation activities. Keywords: transformation, social change, technology assessment, responding to crises, decision-making processes References: Büscher, C. (2018) Framing energy as a sociotechnical problem of control, change, and action. In: Büscher et al. (Eds.): Energy as a Sociotechnical Problem. An Interdisciplinary Perspective on Control, Change, and Action in Energy Transitions. London, New York: Routledge: 14-38. Grunwald, A. (2019): Technology Assessment in Practice and Theory. Routledge Konrad, K.; van Lente, H.; Groves, C.; Selin, C. (2016): Performing and Governing the Future in Science and Technology. In: Felt, U.; Fouché, R.; Miller, C.; Smith-Doerr, L. (Eds.): The Handbook of Science and Technology Studies. 4. Aufl., Cambridge: MIT Press, 465–493. Lösch, A.; Schneider, C.; Dobroć, P.; Frey, P.; Gondolf, J.; Hausstein, A.; Heil, R. (2023): Vision assessment: An orientation framework for the practice of technology assessment 2023. Karlsruher Institut für Technologie (KIT). Schneider, C.; Roßmann, M.; Lösch, A.; Grunwald, A. (2023): Transformative Vision Assessment and 3-D Printing Futures: A New Approach of Technology Assessment to Address Grand Societal Challenges. IEEE transactions on engineering management, 70 (3), 1089–1098. Ufer, U. / Hausstein, A. (2021): „Anthropology of and for Innovation“. In Benoît Godin, Gérald Gaglio, Dominique Vinck (Ed.:): Handbook on Alternative Theories of Innovation, 334–53. Cheltenham: Edward Elgar Publishing

Christine Boshuijzen – van Burken, Deane Baker, Ned Dobos, Milad Ghasri, Erandi Hene Kankanamge, Twan Huybers, Oleksandra Molloy, Jo Plested, Shreyansh Singh	Understanding ethical implications of AI enabled decision support systems on the battlefield	Our future soldiers will increasingly rely on artificial intelligence (AI) enhanced technologies for all sorts of tasks, from reconnaissance to fire support and from countermine applications to medevacs. AI allows systems to perform tasks fully or partially autonomously, meaning that technologies function with no or limited human intervention. In conjunction with the shift from manual and automated technologies to technologies with (semi-)autonomous capabilities on the battlefield, there is the ongoing blurring of distinctions between combatants and non-combatants, between objects or locations that deserve protection under international humanitarian law (such as hospitals and religious spaces) and military objects and what constitutes a normal and what constitutes a diverging pattern of behaviour. AI enabled battlefield decision support systems may be used in support of distinguishing between these categories, in particular vision-based AI support systems. Via real time image recognition, protected symbols (e.g. a Red Cross), gestures (hands up), vehicles or weapons can be brought to the attention of soldiers that are operating on the ground or on the remote end. On the one hand, these systems detect things that human sense perception cannot, increase the amount of information available, and process that information with less bias and more speed than the average human agent. On the other hand, some authors have expressed concerns that reliance on these systems can cause "moral deskilling" (Vallor 2016), whereby human decision-makers gradually lose the ability to deal with ethical dilemmas and tests of integrity that the algorithm cannot deal with. Whichever viewpoint is chosen, it is currently unclear what the actual effects of these technologies are on ethical decision making. Much attention has gone to drafting ethical guidelines for the development of and use of AI in the military [1] , which are laudable attempts to put safeguards in place around AI technologies, but no efforts have been made to test the actual effects of AI on ethical decision making. We pose that ethical guidelines are grounded in (valid) conceptual ideals and theoretical principles, but may benefit from empirical insights to point out where the actual dangers and opportunities are for using AI on the battlefield. There is no actual empirical data on the question of ethical use of AI. Our research project fills this gap. It sheds lights on ethical decision making on the battlefield in an age of AI and autonomous systems. The research question we are interested in is understanding the potential effects of vision AI systems on the battlefield on ethical decision making. We use realistic military shoot/no-shoot decision making scenarios and a commercially available military vision AI system in a pilot study to scope relevant questions to be explored in more detail in future research. This study's contributions involve the analysis of video recordings from a pilot study involving Australian military personnel to understand better the effects of vision-based AI systems on ethical decision-making. This study focuses on hesitation time, shooting judgements, and the influence of AI misclassifications. The results demonstrate that the presence of AI had little impact on shooting judgements in threat scenarios. However, there were significant variations in non-threat scenarios, revealing the potential effects of AI misclassification. Participants that used AI had longer hesitation time before they fired a shot compared to those without AI. References Johnson, J. (2022). Automating the OODA loop in the age of intelligent machines: reaffirming the role of humans in command-and-control decision-making in the digital age, Defence Studies, DOI: 10.1080/14702436.2022.2102486 Santoni de Sio, F., & Mecacci, G. (2021). Four Responsibility Gaps with Artificial Intelligence: Why they Matter and How to Address them. Philosophy & Technology, 34(4), 1057–1084. https://doi.org/10.1007/s13347-021-00450-x Vallor, S. (2016). Technology and the virtues: A philosophical guide to a future worth wanting. Oxford University Press. Verdiesen, I., Santoni de Sio, F., & Dignum, V. (2020). Accountability and Control Over Autonomous Weapon Systems: A Framework for Comprehensive Human Oversight. Minds and Machines. https://doi.org/10.1007/s11023-020-09532-9
Anna Jacyszyn	AI4DiTraRe: how significant and influential Artificial Intelligence is in the Digital Transformation of Research? A multilevel interdisciplinary approach	The recently established Leibniz Science Campus "Digital Transformation of Research" (DiTraRe) investigates the effects of the broadly understood process of digitalisation of research on a multilevel scale. The project concentrates on four research clusters concerning different topics and gathering use cases from varying scientific areas. For a multiscale investigation these research clusters are interwoven with four dimensions, which approach the tasks from different perspectives and pose their own research questions. Within this matrix we are not only developing practical solutions for each use case but also seeking to find generalisations valuable to the scientific community as well as society in general. To gain access to the abundance of information that is available today, search engines and sophisticated information processing applications are required. In order to obtain well-structured knowledge, technologies such as natural language processing, knowledge extraction, and ontology engineering must be applied. Semantic technologies provide a formal representation of knowledge contained in research data, thus facilitating the efficient integration of heterogeneous data sources. The growing adoption of AI-based knowledge mining technologies requires comprehensible and trustworthy AI algorithms ("explainable AI"). Both statistic and linguistic analysis methods as well as machine learning in combination with symbolic logic and interference mechanisms are applied. The aforementioned AI technologies, among others, are investigated within the DiTraRe project by the dimension "Exploration and knowledge organisation". For the use case "Sensitive data in sports science", which is encapsulated in the research cluster "Protected data spaces", our plan is to develop a knowledge graph which will enable sport scientists to easily analyse data and make predictions. In the cluster "Smart data acquisition" we are working with chemists on novel methods of data acquisition, including partial automatisation of this process within the Chemotion Electronic Lab Notebook. The use case "Artificial Intelligence in Biomedical Engineering", enclosed in the cluster "AI-based knowledge realms", will profit from our support concerning introducing large language models into their research. We are cooperating with climate researchers on the use case of "Publication of large datasets" (as a representation of "Publication cultures" cluster) where we are employing AI techniques for an unchallenging organisation of very large amounts of research data.