fPET 2023

2023 Forum on Philosophy, Engineering, & Technology

19-21 April 2023
Program Booklet

fPET 2023 is proudly supported by:

European Research Council
Established by the European Commission

TU Delft

threejoy

BIGBEACON

Organization

Conference co-chairs
Ibo van de Poel (Delft University of Technology)
David Goldberg (ThreeJoy Associates, Inc.)

Program chair
Neelke Doorn (Delft University of Technology)

Local organizing committee
Lotte Asveld (Delft University of Technology)
Michael Klenk (Delft University of Technology)
Nathalie van den Heuvel (Delft University of Technology)
Freek van der Weij (Delft University of Technology)
Monica Natanael (Delft University of Technology)

Program Committee
Sabine Ammon (TU Berlin)
Stephen C. Armstrong (University of Toronto)
Christine Boshuizen - van Burken (UNSW Sydney)
Anders Buch (VIA University College)
Steen Hyldgaard Christensen (Aalborg University)
Rockwell F. Clancy (Virginia Tech)
Benjamin Cohen (Lafayette College)
Shannon Conley (James Madison University)
Darryl Cressman (Maastricht University)
Christelle Didier (Université Charles de Gaulle -Lille 3)
Richard Evans (Cornell University)
Darryl Farber (PennState)
Rider Foley (University of Virginia)
Albrecht Fritzsche (Rabat Business School, Morocco)
Dennis Gedge (Consulting engineer)
Matthew Harsh (California Polytechnic State University)
Joseph Heckert (North Carolina State University)
Justin Hess (Purdue University)
Deborah Johnson (University of Virginia)
Diane Michelfelder (Macalester College)
Glenn Miller (Texas A&M)
Byron Newberry (Baylor University)
Diana I. Pérez (Universidad de Buenos Aires)
Zachary Pirtle (NASA and George Washington University)
Michael Poznic (KIT Karlsruhe)
Édison Renato Pereira da Silva (Universidade Federal do Rio de Janeiro)
Joe Pitt (Virginia Tech)
ZHU Qin (Virginia Tech)
Andrés Santa-María (Universidad Técnica Federico Santa María)
Viola Schiaffonati (Politecnico di Milano)
Steven Umbrello (Institute for Ethics and Emerging Technologies)
Pieter Vermaas (TU Delft)
Bruce Vojak (University of Chicago)
Mark Thomas Young (University of Vienna)
Sjoerd Zwart (TU Delft)

Practical Information

Travel Information

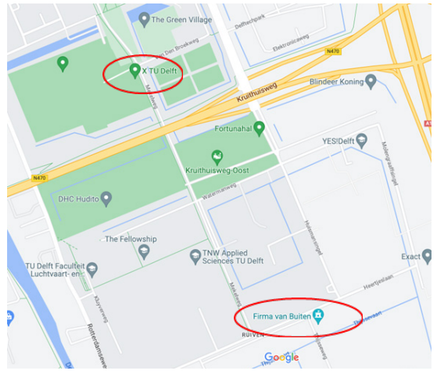
General travel information
fPET2023 will take place in building X (Mekelweg 8, 2628 CD Delft), on the TU Delft campus. The venue is easily reachable by bus from the train station Delft. You can take bus 40 (in the direction of Rotterdam Centraal), 69 (in the direction of TU Campus) or 174 (in the direction of Rotterdam Noord) to bus stop 'Berlageweg'. In the morning of Wednesday 19th, you can for example take:

- Bus 40 at 08:18
- Bus 174 at 08:23
- Bus 69 at 08:30
- Bus 40 at 08:38
- Bus 69 at 08:45

The dinner on Thursday evening is the only part of the conference that is not in building X. It will instead be at Firma van Buiten (Thijsseweg 1, 2629 JA Delft). You can walk (10-15 min) or take a taxi to the restaurant. From Firma van Buiten to the city centre, we have arranged a shuttle bus driving up and down to the train station. The first departure time is 20:45. You can of course also take a taxi on your own initiative. By foot, getting to the city center will take about 40 minutes. On Friday, there is room in X to leave your luggage if you checked out at your hotel that morning. When leaving X on Friday, you can take bus 40, 69, or 174, all in the direction of the station. They leave every 5/10 minutes.

Travelling by train from Schiphol to Delft
Please note that there are currently no direct trains from Schiphol to Delft due to an earlier accident. Instead, it is probably best to take the high speed train from Schiphol directly to Rotterdam and from there to take a train to Delft. With a valid train ticket to Delft, you can board any train from Rotterdam to Delft. Note however, that you need to pay a supplement for the high speed train. Tickets and the supplement can be bought at Schiphol from the NS train ticket machines. Actual train times can be found at https://www.ns.nl/en

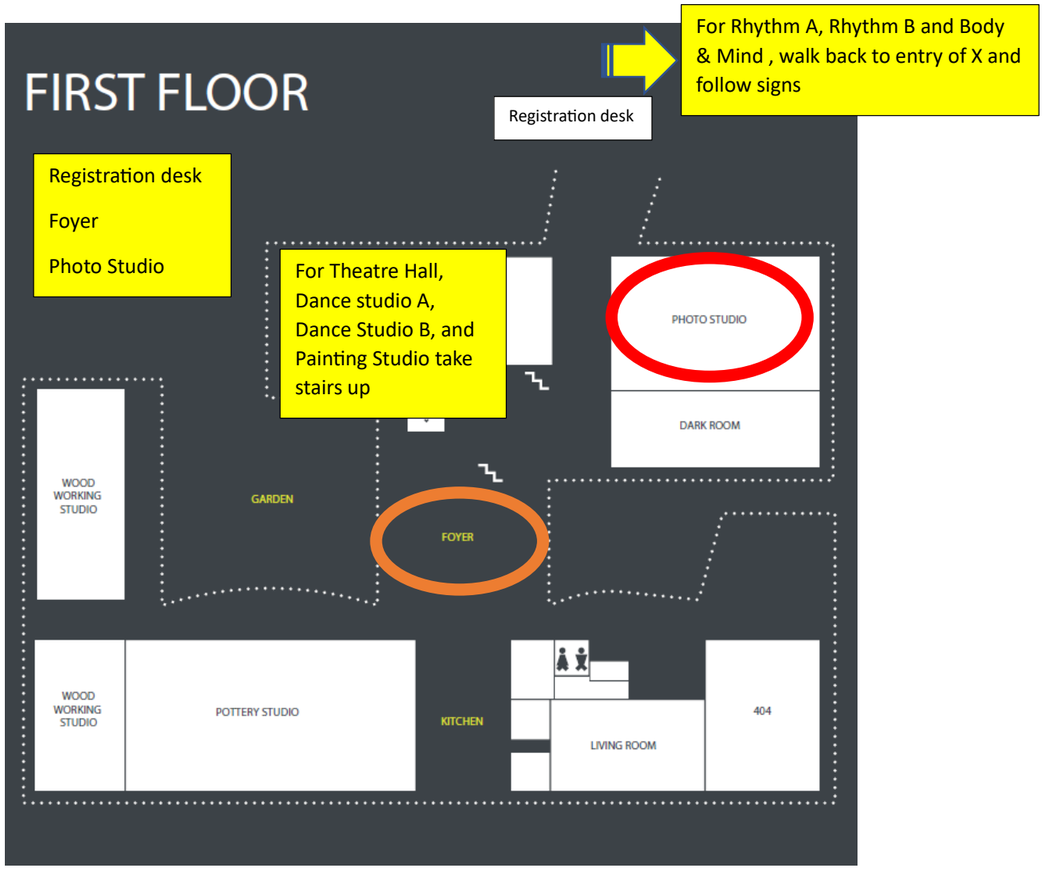
Travelling by public transport in the Netherlands
The Netherlands has a quite good public transport network. To plan trips and buy e-tickets you can use this website: https://9292.nl/en
If you wish you can also download the 9292 app on your phone; just search for 9292 in the app store.
You can also use public transport to travel from your hotel to the conference location. For the conference location use 'Mekelweg 8 Delft'
After the dinner on Thursday, we will have coaches that bring you back to Delft railway station in the center of Delft.

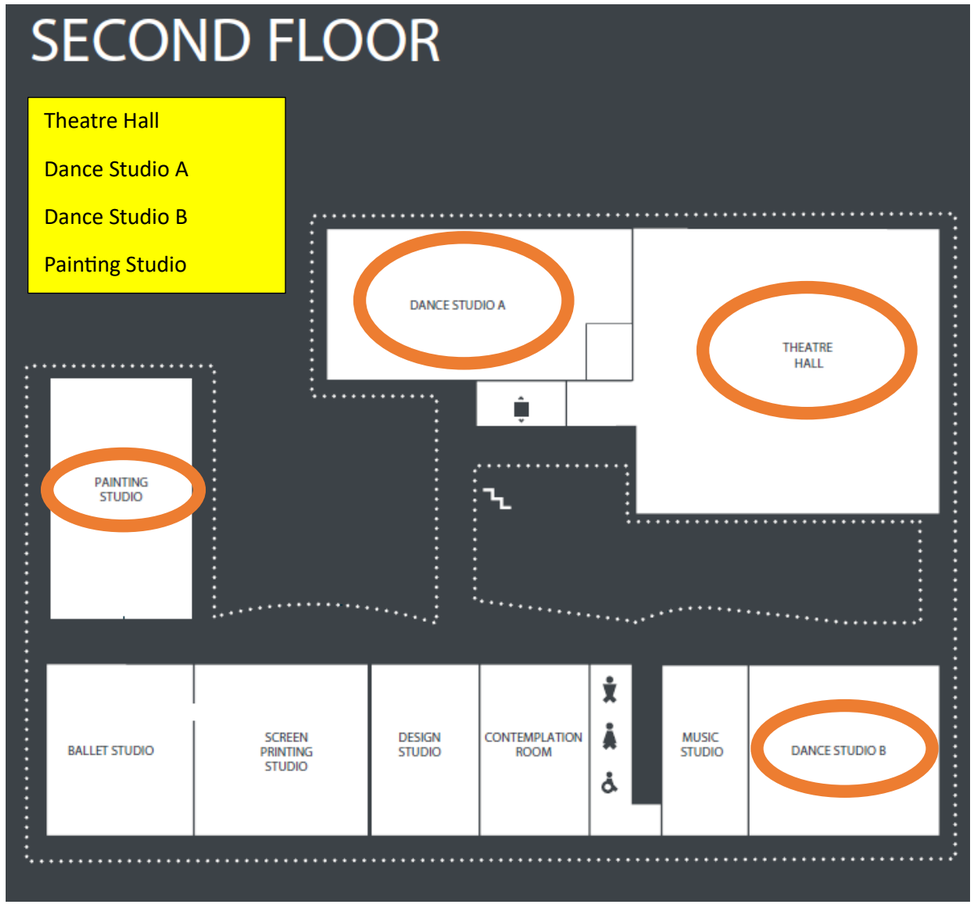


Information on presentations

The presentations are grouped in sessions that contain 2 or 3 papers. Every presentation should take a maximum of 20 minutes so 10 minutes remain for the Q&A. A chair is appointed for every session who makes sure presenters stick to these times. The presentations are in person only.

Information on building/rooms





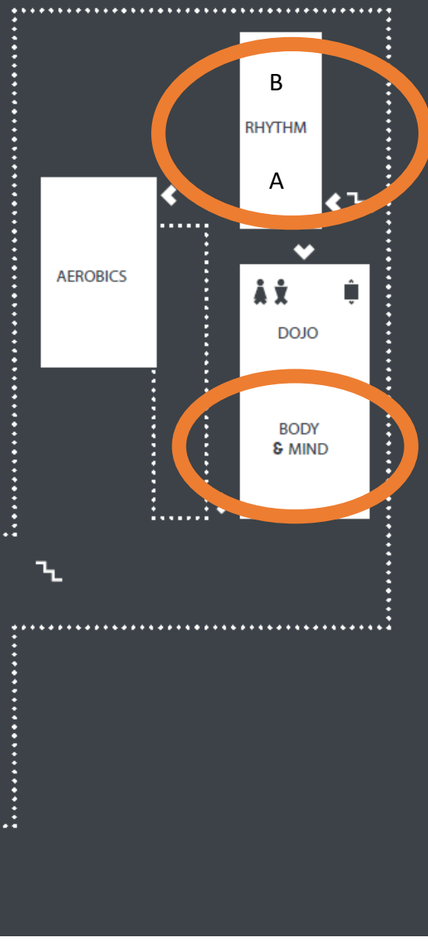
For Rhythm A, Rhythm B and Body & Mind, walk back to entry of the X building and follow signs

(from entry straight on, then turn right, and take stairs to first floor)

Rhythm A

Rhythm B

Body & Mind



Theatre Hall

Dance Studio A

Dance Studio B

Painting Studio

First floor not connected, walk back to entry at ground floor instead

Rhythm A

Rhythm B

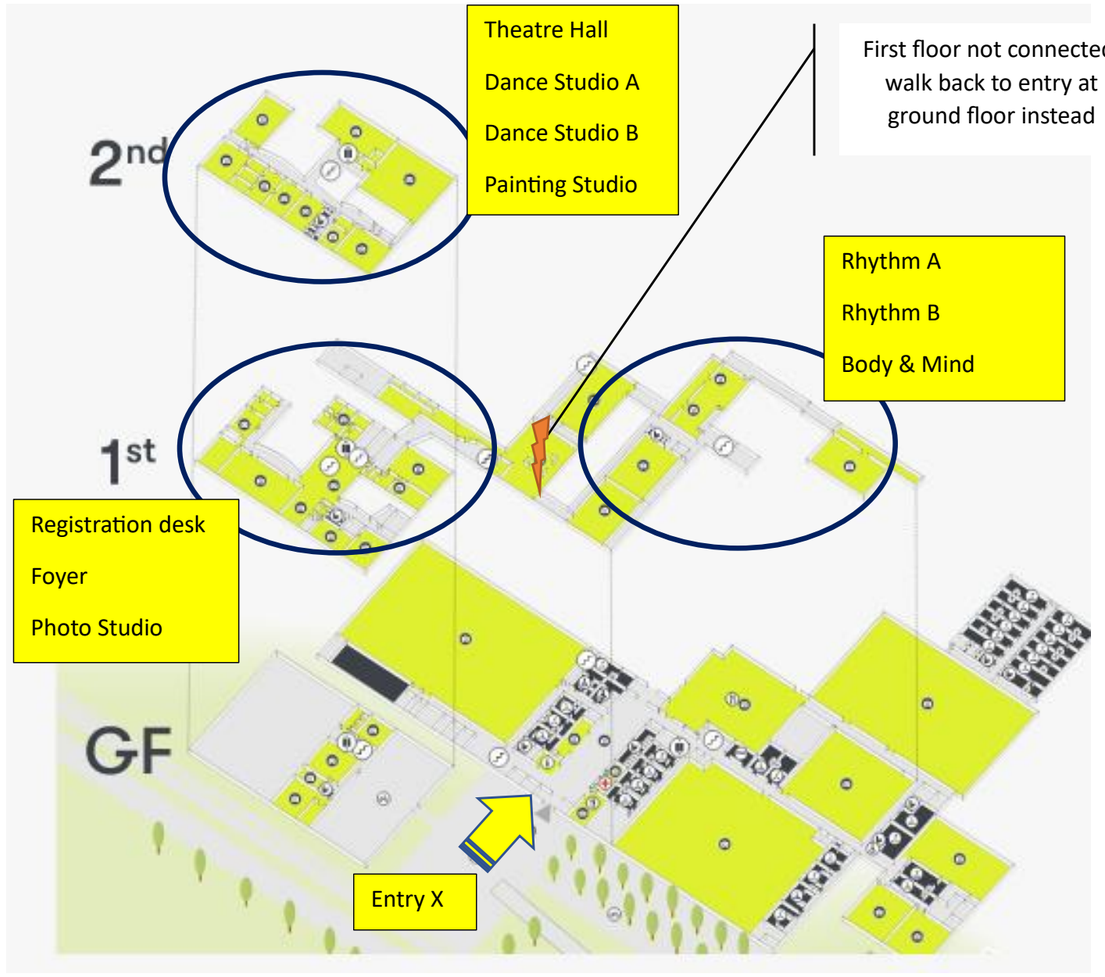
Body & Mind

Registration desk

Foyer

Photo Studio

Entry X



Program fPET 2023

The entire fPET program will take place at TU Delft, in building X, except for the dinner on Thursday, which is at Firma van Buiten.

Wednesday, April 19th			
Time	**Session**	**Content**	**Location**
08:30-09:00	**Registration**		Foyer
09:00-09:30	**Welcome** David Goldberg and Ibo van de Poel		Theatre Hall
09:30-10:30	**<u>Session 1:</u>** Ibo van de Poel. Technology, design and value change Chair: Lotte Asveld		Theatre Hall
10:30-11:30	**<u>Session 2A:</u> Autonomous vehicles**. *Chair: Sebastian Krügel*	▪ Michael W. Schmidt. Ethical Issues Concerning the International Trade of Autonomous Driving Technology ▪ Sebastian Krügel & Matthias Uhl. The risk ethics of autonomous vehicles: A continuous trolley problem in regular road traffic	Body & Mind
	<u>Session 2B:</u> Technical Functions. *Chair: Clint Hurshman*	▪ Ryan Wittingslow. Designed Things, Technical Functions, and Speech Act Theory ▪ Clint Hurshman. A Taxonomy for Artifact Function Pluralism	Rhythm A
	<u>Session 2C:</u> Genetic modification for health. *Chair: Lieke Baas*	▪ Mattia Pozzebon. Animal Models and Human Diseases: The Ethics of Genome Editing Animals to Cure Humans ▪ Lieke Baas, Rieke van der Graaf, Karina Meijer & Annelien Bredenoord, for the SYMPHONY consortium. Techno-fixing hemophilia away: an ethical assessment of arguments for gene therapy	Rhythm B
	<u>Session 2D:</u> Language and communication. *Chair: Martin Stacey*	▪ Rick Evans. Bringing back words from their metaphysical to their everyday use: How Wittgenstein and cognitive-functional linguistics can help us to better understand and teach engineering communicative practice. ▪ Martin Stacey & Claudia Eckert. Epistemological challenges for multidisciplinary design	Theatre Hall
	<u>Session 2E:</u> Engineering ethics in global perspective. *Chair: Behnam Taebi*	▪ Patricio Quintana Gallo & Carlos Verdugo Serna. The Rise and Fall of the Code of Ethics of the Chilean Engineering Council: History and Vicissitudes ▪ Ali Dizani, Behnam Taebi, Ibo van de Poel & Amineh Ghorbani. A comparative institutional analysis of engineering ethics in the Netherlands and the United States	Dance Studio B
	<u>Session 2F:</u> Global governance of biotechnology. *Chair: Lotte Asveld*	▪ Adam McCarthy. Knowledge-policy interfaces for wicked problems. The Co-production of Digital Sequence Information governance at the Convention for Biological Diversity. ▪ Bob Kreiken & Lotte Asveld. How to develop principles for fair access and benefits of Digital Sequence Information for scientists and engineers	Photo Studio

11:30-12:30	**Lunch**		Foyer
12:30-14:00	**<u>Session 3A:</u> Panel Ecosystem-education**	▪ Panel Ecosystem-education and universities' social responsibility. Organizers: Gunter Bombaerts, Diana A. Martin, Neelke Doorn, Zoë Robaey, Jeremy Mantingh & Suleman Audu.	Body & Mind
	<u>Session 3B:</u> (Post)phenomenology. *Chair: Lars Botin*	▪ Bouke van Balen. Why We Need a Phenomenology of Brain Computer Interface-mediated Communication ▪ Lars Botin & Tom Børsen. Denominating the Movements in the Lemniscate of Technology	Photo Studio
	<u>Session 3C:</u> Just and Fair AI. *Chair: Verity Truelove*	▪ Bauke Wielinga. Complex Fairness ▪ Florian Richter. Justice and Smart Societies – Algorithmizing Fairness ▪ Verity Truelove, Oscar Oviedo-Trespalacios, Kayla Stefanidis & Levi Anderson. Differences in perceptions of justice between police and AI-based automated enforcement	Theatre Hall
	<u>Session 3D:</u> Social Media and public governance. *Chair: Andreas Spahn*	▪ Paige Benton & Michael W. Schmidt. The Harm of Social Media to Public Reason ▪ Andreas Spahn. Conflict, Compromise or Consensus? - Social Media and Democracy ▪ Irene Olivero. A Pragmatic Method for Engineering Concepts and Technologies	Rhythm A
	<u>Session 3E:</u> Value change. *Chair: Simona Aracri*	▪ Yunxuan Miao. Understanding sustainability as a changing value in lithium battery technologies ▪ Karen Moesker. Direct Potable Water Reuse – Change inducing for moral values? ▪ Simona Aracri, Rosangela Barcaro & Massimo Caccia. Women in the Ever Evolving Marine Robotics	Rhythm B
14:00-14:30	**Break**		Foyer
14:30-16:00	**<u>Session 4A:</u> Workshop Designing philosophies of education**	▪ Workshop Designing philosophies of education as fruitful bridges between philosophical ideas of engineering and realities in engineering education. Organizer: Daniel Marom.	Body & Mind
	<u>Session 4B:</u> Models and Epistemology. *Chair: Francisco Olivera*	▪ Zachary Pirtle. How do engineers model engineering work? Contrasting cost and schedule models to understand modeling practices in complex system developments ▪ Vivek Kant & Michael Poznic. Epistemic achievements of engineers in relation to sociotechnical systems: From technological knowledge to engineering understanding ▪ Franscisco Olivera & Martin Peterson. Epistemic uncertainties in flood frequency analyses and the ethics of belief	Photo Studio
	<u>Session 4C:</u> Ethics of Technology. *Chair: Fernando Secomandi*	▪ Brandon Long. A Frankian Problem with Genetic Engineering ▪ Camilla Quaresmini, Eugenia Villa, Valentina Breschi, Viola Schiaffonati & Mara Tanelli. Qualification and quantification of fairness for diversity-aware and inclusive mobility ▪ Fernando Secomandi. Towards a postphenomenology of mutual recognition in AI design	Dance Studio B
	<u>Session 4D:</u> AI Ethics. *Chair: Michael Klenk*	▪ Sara Blanco. The Normative Need for Trust in AI ▪ Jens Pilger & Michael W. Schmidt. AI as Inventor? A patent system's call for philosophy ▪ Michael Klenk. Manipulative technology	Theatre Hall

	<u>Session 4E:</u> Goals of Engineering Education. *Chair: Anders Buch*	▪ Lavinia Marin. What should we strive for in engineering ethics education? A complex and granular mapping of Engineering Ethics Education competencies and skills ▪ Katherine A. Goodman. Metaphors of Learning and Their Uses in Engineering Education: Reflections from an Engineering Educator ▪ Anders Buch & Steen Hyldgaard Christensen. Rethinking Engineering Professionalism through the Concept of Bildung	Rhythm A
	<u>Session 4F:</u> Value Change. *Chair: Diana Martin*	▪ Anthony Longo. Revaluing the Public Debate: A Postphenomenological Account of Changing Values in the Digital Public Sphere ▪ Christine Boshuijzen-van Burken, Shannon Spruit, Lotte Fillerup & Niek Mouter. Using Participatory Value Evaluation to reach unorganised affected stakeholders of military autonomous systems ▪ Diana Martin, Rockwell F. Clancy, Qin Zhu & Gunter Bombaerts. Why do we need Norm Sensitive Design? A WEIRD critique of value sensitive approaches to design	Rhythm B
16:00-17:00	**<u>Session 5:</u>** Keynote Guru Madhavan. The Politics—and Philosophy—of Prestige in Engineering. Chair: Neelke Doorn		Theatre Hall
17:00-17:30	**Break**		Foyer
17:30-19:00	**<u>Session 5A:</u> Workshop Value Change in AI and Robot Systems**	▪ Workshop AI, Robots and Value Change: Exploring Anthropomorphism's Normative Impact on Day-to-Day Social Practices. Organizers: Tom Coggins & Olya Kudina.	Photo Studio
	<u>Session 5B:</u> Workshop Value Change in Energy Systems	▪ Workshop Visions of Change. Battling perfect moral storms in the context of energy transition. Organizers: Anna Melnyk & Joost Alleblas.	Rhythm A
	<u>Session 5C:</u> Workshop Value Change and Climate Engineering	▪ Workshop Techno-Moral Change and the Role of Scenario Thinking and Modelling in Solar Climate Engineering Research. Organizers: Benjamin Hofbauer, Behnam Taebi, Giacomo Marangoni & Gideon Futerman. .	Rhythm B
	<u>Session 5D:</u> Poster session	▪ For the list of posters that will be on display, please see the bottom of this document.	Dance Studio A
19:00	**Reception**		Foyer

Thursday, April 20th			
Time	**Session**	**Content**	**Location**
09:00-17:00		The posters are exhibited all day in Dance Studio A	Dance Studio A

	<u>Session 6A:</u> Design for Values. *Chair: Lotte Asveld*	▪ Marula Tsagkari, Mar Pérez-Fortes & Ibo van de Poel. A Value Sensitive Design (VSD) approach for the sustainable design of multiscale CO2 electrochemical conversion (CO2ER) ▪ Sabine Ammon, Nele Fischer, Tim Hildebrandt, Steffen Müller, Dieter Peitsch & Utz von Wagner. Value Change by Ethical Vision Design: Integrated ethics methods for developing the research agenda in an interdisciplinary engineering project – A field report from the research network PureMobility ▪ Britte Bouchaut, Thom Tribble & Lotte Asveld. Safe-by-Design: The need for engineered prerequisites for inherent safety.	Body & Mind
	<u>Session 6B:</u> Decolonial perspectives. *Chair: Andrea Gammon*	▪ Dimpho Radebe & Kai Zhuang. Confronting the shadows of colonialism in engineering education: An ongoing practice of "Calling In" ▪ Joseph Emmanuel D. Sta. Maria & Matthew J. Dennis. Rethinking McMindfulness from an East Asian Perspective ▪ Andrea Gammon. The "waiting room version of history" in post-colonial- and sustainable-engineering projects	Theatre Hall
09:00-10:30	**<u>Session 6C:</u> Political Philosophy and Technology.** *Chair: Avigail Ferdman*	▪ Maarten Franssen. The ethics of engineering vs. the politics of technology ▪ Jérémie Supiot & Philippe Estival. How democratic debates contribute to a better governance of innovation ▪ Avigail Ferdman. Perfectionism: Towards a political philosophy of technology and human flourishing	Dance Studio B
	<u>Session 6D:</u> AI and Epistemology. *Chair: Ekaterina Bogdanova*	▪ Dmitry Muravyov. Datafied doubt as an epistemic virtue: learning from activists' and practitioners' perspectives ▪ Ekaterina Bogdanova. Mind-body-data: Epistemic shift of digital psychiatry ▪ Caitlin Grady & Lauren Dennis. Ethical-epistemic values driving decisions under uncertainty: The case of hydropower in the United States	Painting Studio
	<u>Session 6E:</u> Value Change. *Chair: Tomasz Żuradzki*	▪ Jon Rueda. Reprogenetic technologies, future value change, and the axiological underpinnings of reproductive choice ▪ Rafael Coimbra & Édison Renato Pereira da Silva. How to represent human values in the Metaverse ▪ Tomasz Żuradzki. Genetic enhancement in bioethical discussions: a computational approach to value change	Photo Studio
10:30-11:00	**Break**		Foyer
11:00-12:30	**<u>Session 7A:</u> Workshop Education of a Whole New Engineer**	▪ Workshop An UnWorkshop for the Education of a Whole New Engineer. Organizers: David E. Goldberg, John R. Donald, Beata Francis & Katherine A. Goodman.	Body & Mind
	<u>Session 7B:</u> Philosophy of Technology. *Chair: Mark Ryan*	▪ Pietro Camin & Samuela Marchiori. Are ethics guidelines enough? On the effectiveness of ethics guidelines for the design, development, and deployment of conversational agents ▪ Agostino Cera. The Ontophobic Turn (Between Philosophy of Technology and Engineering)	Theatre Hall

		▪ Mark Ryan. A Critique of Human-Centred AI: Foucault, Power, and Philosophical Anti-humanism	
	<u>Session 7C:</u> Engineering Education. *Chair: Rider Foley*	▪ Katherine Brichacek, Ordel Brown & Laura Pigozzi. Explicit empathy instruction at the intersection of engineering design, technical communication, and philosophy ▪ Andrea Gammon, Qin Zhu, Rockwell Clancy, Scott Streiner & Ryan Thorpe. Critically Examining the Broader Implications of Methodological Design in Cross-Cultural, Multi-sited Case Studies of Engineering Ethics Education ▪ Rider Foley. Tacking against the headwinds of the military-industrial complex: Engineer's values and professional formation	Dance Studio B
	<u>Session 7D:</u> Value Change. *Chair: Gunter Bombaerts*	▪ Michael J. Bernstein, Lauren Withycombe Keeler, Luke Boyle & John Harlow. A needs-based theory of social value for research and innovation ▪ Michael Klenk. The Gap in Value Change ▪ Gunter Bombaerts. Adding to understanding value change with Luhmann's event-concept	Photo Studio
12:30-14:00	**Lunch**		Foyer
12:30-14:00	**Poster Session**		Dance Studio A
14:00-15:00	**<u>Session 8:</u>** Keynote Sarah Spiekermann: On the nature of values and their role in IT system design Chair: Olya Kudina		Theatre Hall
15:00-15:30	**Break**		Foyer
15:30-17:00	**<u>Session 9A:</u> Interdisciplinary studies of technology and engineering.** *Chair: Artur Bogucki*	▪ Syafira Fitri Auliya. Adversarial Machine Learning for Deliberating Decision-Making in Elections ▪ Aurélien Béranger & Hugues Choplin. Autonomy in the test of low-tech devices. ▪ Artur Bogucki & Paula Gürtler. Ethical, Legal, and Socio-Economic Analysis of Agrifood Data Marketplaces - Lessons from FlexiGroBots	Body & Mind
	<u>Session 9B:</u> AI and Epistemology. *Chair: David Spurrett*	▪ Daniele Chiffi, Mattia Petrolo, Viola Schiaffonati & Giacomo Zanotti. Beyond reliability: values and Trustworthy AI ▪ Stefano Canali, Viola Schiaffonati & Andrea Aliverti. Data Quality as an Emerging Issue for AI: The Case of Wearable Technology ▪ David Spurrett. Manipulating the scaffolded user	Theatre Hall
	<u>Session 9C:</u> Moral responsibility in engineering. *Chair: Erhardt Graeff*	▪ Zoë Robaey. A virtue ethics account of moral responsibility ▪ David M. Douglas & Justine Lacey. Defining Ethical Risk for Technology ▪ Erhardt Graeff. Using Civic Professionalism to Frame Ethical and Social Responsibility in Engineering	Dance Studio B

	<u>Session 9D:</u> Philosophy of human – machine relations. *Chair: Olya Kudina*	▪ Rafael Mestre & Anibal Monasterio. Biorobots as objects, tools or companions? An ethical approach to understand bio-hybrid systems ▪ Llona Kavege, Julia Hermann & Olya Kudina. Machine Mediated Gestation - Ontological Distinctiveness and Normative Implications	Painting Studio
	<u>Session 9E:</u> Value change. *Chair: Milenko Budimir.*	▪ Hugo Paris, Nicolas Freud & Caroline Ladage. Training engineers for sustainability, but which one? A discussion on alternatives to the “good Anthropocene” ▪ Antoine Bouzin. The Critique of Technology by French Environmentalist Engineers: a Shift in Values ▪ Milenko Budimir. The role of “honor” in engineering a solution to climate change	Photo Studio
17:00	**End of program**		
17:30	**Gather to walk (+- 15 min) to Firma van Buiten**		Registration Desk
18:00-21:00	**Dinner**		Firma van Buiten Thijsseweg 1, 2629 JA Delft

Friday, April 21st			
Time	**Session**	**Content**	**Location**
09:00-10:30	**<u>Session 10A:</u> Engineering in the context of climate risk and uncertainty.** *Chair: Frieder Bögner*	▪ Gideon Futerman. What does a "Science" of Existential Risk look like? The future of Existential Risk Studies ▪ Frieder Bögner. The lock-and-key model: Opportunities and challenges for technosolutionism with regard to artificial intelligence.	Dance Studio A
	<u>Session 10B:</u> Applied (post)phenomenology and technological mediation. *Chair: Andrej Dameski*	▪ Dylan Wittkower. Understanding Designing to Match Users’ Mental Models as Eidetic Reduction ▪ Hans Voordijk, Seirgei Miller & Faridaddin Vahdatikhaki. Operator guidance systems in road construction: a technological mediation perspective ▪ Andrej Dameski, Andreas Spahn, Alessandro Corbetta, Antal Haans, Caspar A. S. Pouw, Federico Toschi, Jaap Ham, Rabia I. Kodapanakkal & Gunter Bombaerts. Empirics of collectives: using empirical data from Dutch train stations to study collective formation, behaviour, and use of technology	Rhythm A
	<u>Session 10C:</u> Value Change. *Chair: Byron Newberry*	▪ Samuela Marchiori. Technology-mediated moral change. Exploring the relation between moral concept change (MCC) and moral value change (MVC) ▪ Aznavur Dustmamatov. The Role of Artifacts in Value-Formation: From Affordance to Exposure ▪ Byron Newberry. The Dynamics of Technological Value-Ladenness: A Pragmatic Approach	Theatre Hall
	<u>Session 10D:</u> Panel Maintenance &	Panel Maintenance & Philosophy of Engineering: ▪ Mark Thomas Young - Evolving Infrastructure: Maintenance and the Political Lives of Artifacts	Photo Studio

	Philosophy of Engineering.	▪ Rebecca Mossop - Communicating Smoothly: How Material Breakdown is not the only thing to be Maintained in a Large Technical System ▪ Alexandra Karakas - How can Science and Technology Studies benefit from the Philosophy of Maintenance? ▪ Steffen Steinert - Making Things Durable: Maintenance of Technology as Maintenance of Society and Value	
	<u>Session 10E:</u> Value Change Workshop	▪ Workshop Value Change and Engineering Organizers: Ibo van de Poel & Freek van der Weij.	Rhythm B
10:30-11:00	**Break**		Foyer
11:00-12:30	**<u>Session 11A:</u> New methods in engineering education.** *Chair: Christelle Didier*	▪ Kai Zhuang, Dimpho Radebe & Mojgan Jadidi. Art-Inspired Pedagogies in Engineering Education - Using Comics, VR/AR, Gaming, and Music in Engineering Education ▪ Elin Sporrong, Cormac McGrath & Teresa Cerratto Pargman. Situating values in practice: University teachers' perspectives on assessment and emerging AI ▪ Christelle Didier. Rethinking engineering education in a French-speaking context using the concept of Bildung	Dance Studio A
	<u>Session 11B:</u> Natural sciences and conceptual change/engineering. *Chair: Terry Bristol*	▪ Alok Srivastava. The Development of the Affordances of Indicator Diagrams into Thermodynamic Work Cycles in the Material and Theoretical Cultures of Early Steam Engines ▪ Terry Bristol. Christiaan Huygens and the Foundations of the Engineering Thermodynamic Worldview	Photo Studio
	<u>Session 11C:</u> Practitioners' reflections. *Chair: Zachary Pirtle*	▪ Dennis Gedge. The individual and Civil Engineering ▪ Marc Steen, Jurriaan van Diggelen, Friso van Houdt, Robin van den Akker & Leanne Cochrane. Practitioners' reflections on organizing Transdisciplinary Innovation ▪ Zachary Pirtle, Claudia Eckert, Rémi Gandoin, Nina Jirouskova, Michael Poznic, Beth-Anne Schuelke-Leech, Martin Stacey & Loretta von der Tann. Knowledge for Practicing Engineers and Philosophers: Reconsidering Walter Vincenti's What Engineers Know and How They Know It	Rhythm A
	<u>Session 11D:</u> Panel Responsibility in Practice	▪ Panel Responsibility in Practice. Organizer: Zoë Robaey.	Rhythm B
	<u>Session 11E:</u> Value change workshop	▪ Workshop Value Change and Engineering. Organizers: Ibo van de Poel & Freek van der Weij.	Theatre Hall
12:30-13:30	**Lunch**		Foyer
12:30-13:30	**Business meeting for steering committee** Only for members of the fPET steering committee		Photo Studio
13:30-14:30	**<u>Session 12:</u>** Keynote Samantha Kleinberg – From Causes to Actions Chair: Martin Sand		Theatre Hall

14:30-16:00	**<u>Session 13A:</u> Ethics of Engineering Education.** *Chair: Sabine Ammon*	▪ Lukas Fuchs & Gunter Bombaerts. The Limits of Co-creation? The Ethics of Industry-funded Student Teams at Engineering Universities ▪ Sabine Ammon, Alexandra Kljagin, Juliane Rettschlag & Martina Vortel. Value Change by Value Elicitation, Value Amplification, and Value Integration in Interdisciplinary Ethics Education: Examples from the Berlin Ethics Certificate	Dance Studio A
	<u>Session 13B:</u> Codes of conduct *Chair: Diane P. Michelfelder*	▪ Pieter Vermaas. Toward engineering codes of conduct for philosophy of technology ▪ Diane P. Michelfelder & Sharon A. Jones. Engineering codes of ethics with care	Rhythm A
	<u>Session 13C:</u> Environmental and energy ethics. *Chair: James Hutton*	▪ Cristina Richie. The Environmental "Paradox of Prevention" in Socially Disruptive Medical Technologies ▪ James Hutton. Turn, Turn, Turn: Wind Turbines, Aesthetic Persuasion, and Changing Feelings	Theatre Hall
	<u>Session 13D:</u> Panel Designing with nature	Panel Designing with nature: reflections on biomimetic practices. • Alessio Gerola, Zoe Robaey, Vincent Blok: What does it mean to mimic nature? A typology for biomimetic design • Dayo Jansen: Understanding animal flight and escape to create and improve bio-inspired robots • Jaco Appelman: Bio-inspired Innovation • Paul Breedveld: BITE – Bio-Inspired Technology Group	Photo Studio
	<u>Session 13E:</u> Value change. *Chair: Oyku Ulusoy*	▪ Franziska Poszler & Benjamin Lange. AI-enabled decision-support systems: A catalyst or curse for human's ethical decision-making? ▪ Sarah E. Carter & Heike Felzmann. How do we value data privacy? Initial results from semi-structured interviews ▪ Oyku Ulusoy. The Interconnection of Subjective Values and Experience-Based Learning	Rhythm B
16:00-17:00	**Closing** Chair: Ibo van de Poel Discussion of fPET future and bylaws Springer's book series Philosophy of Engineering and Technology (Pieter Vermaas).		Theatre Hall

Poster Session	
Author(s)	**Title**
Maximilian Rossmann	The ChatGPT hype as a medium to popularize and jeopardize social debates about values and virtues in education.

Belen Liedo	Roboticizing Emotions? A Framework for Analyzing Care Workers' Emotions towards Assistive Robots.
Antonia Kempkens	A "Data Trustee" as Mediator between the Value of Privacy and Data Analysis
Olaya Fernández Guerrero, Ángel Luis Rubio García & Julio Rubio García	Virtue Ethics and software development: debates and proposals
Tilke Devriese	Tactics & Rhetorics. Images of calculation models in architectural design competitions.
Rockwell F. Clancy, Ingvild Bode & Qin Zhu	Weaponized artificial intelligence (AI), normative and culture psychology: Methodological contributions to policies on emerging technologies
Martin Sand, Michael Klenk & Andrea Gammon	Assessing student progress in engineering ethics education
Franziska Poszler, Edy Portmann & Christoph Lütge	Translating ethical principles into computer codes: Experts' opinions on why (not) and how to do it
Alik Pelman	Technological Optimism vs. Scepticism: Agroecology as a Case Study
Joshua Wodak	Towards A Philosophy of Bioengineering and Geoengineering: Risk Ethics for a Climate Change Technofix
Rockwell F. Clancy, Qin Zhu, Louis Hickman, Subhabrata Majumdar, David McGraw & Andrew Katz	Towards a global ethics of artificial intelligence (AI): Ensuring AI ethics is culturally responsive and psychologically realist
Nynke van Uffelen	Introducing the concept of hidden morality in energy justice
Michael Wartmann	Designing Automation with Care
Eswaran Subrhamanian, Anne-Françoise Schmid, Ira Monarch, Muriel Mambrini-Doudet & Anne-Lise Dauphiné-Morer	Generic design Science and lived experiences
Sebastian Krügel, Andreas Ostermaier & Matthias Uhl	The moral authority of ChatGPT
Gideon Futerman	Scenarios for the development of SRM between 2030-2050 with particular reference to the interaction between SRM and Climate Change and Global Catastrophic and Existential Risk
Udo Pesch	Automation and autonomy: How sociotechnical systems disrupt the counterfactual status of autonomy
Emmanuel Caillaud	Ethics and sustainability concepts for the factory of the future: what should we teach and how?

fPET 2023 is organized in cooperation with the project 'Design for Changing Values', which has been funded as Grant Agreement No788321 by:

European Research Council
Established by the European Commission

Abstracts (in chronological order)

Wednesday

Session 1: Technology, design and value change

Ibo van de Poel

Values play an important role in technological development and design. First, they are one of the drivers of technological choices and design. Second, they are important for the social acceptance of newly designed technologies, i.e., they play a role in whether users and stakeholders are willing to accept a technology, and use it. Third, they are key in judging the moral acceptability of technology and design choices.
There are now several approaches, such as design for values and responsible innovation, that aim at pro-actively addressing values in technological design, in order to increase both social acceptance and moral acceptability. Such approaches tend to assume that values are static, i.e., that they do not change after a technology has designed. However, it seems that values can change over time, and that we need approaches that can address such value change.
When it comes to value change, it is important to make a distinction between changes in valuing and changes in what is valuable. It seems obvious that there are changes in what people have valued over time. Such changes in valuing may alternatively be described as changes in descriptive values and they seem particularly important for the social acceptance of technology. When it comes to moral acceptability, however, we are interested in what is valuable form a normative, or moral point, of view. It is less obvious that such normative values can also change.
I will suggest that philosophical pragmatism, particularly John Dewey's writings about value, offer a way to understand the possibility of normative value change. Pragmatists may understand values as 'evaluative devices' that are functional in recognizing, judging and ultimately resolving moral problems. On such an understanding, values may change if and when we are confronted with new types of moral problems. Moreover, what values are normatively justified does not depend on what people actually desire or value, but primarily on whether interventions based on such values are ultimately able to resolve morally problematic situations. I will illustrate the possibility of such value change by a number of examples in the domain of technology.
In the last part of my talk, I will discuss how engineers and designers can better address the possibility of value change in design processes. I will suggest that there are roughly three strategies for doing so, namely 1) trying to anticipate possible value changes already during the design phase of new technology, 2) morally experimenting with new technology and monitoring value change over time, 3) following a number of design principles that allow adjusting technological designs over time so allowing the possibility to adjust designs to new values.
If you are interested, you can download the white paper on value change and technological design here.

Session 2A: Ethical Issues Concerning the International Trade of Autonomous Driving Technology

Michael W. Schmidt

This paper focuses on a rather novel ethical problem regarding autonomous driving technology (ADT): How should liberal democracies deal with the international trade of ADT?

Extensive ethical research has uncovered a plethora of issues concerning ADT. These include, for example, issues of automated decision-making and risk distribution in crash scenarios and mundane driving situations, privacy issues, and issues regarding moral responsibility and legal liability. Since the use of ADT can have an impact on human or basic rights, many of these issues call for moral solutions in the political realm and corresponding regulation (Brändle & Schmidt, 2021; Himmelreich, 2018; Rodríguez-Alcázar et al., 2020). However, even if there is a specific solution to such an ethical issue of ADT that is justified within a specific liberal democracy, it is possible that this solution is unacceptable for a different liberal democracy.

This leads to ethical issues regarding import and export of ADT between liberal democracies. Think of the classical problem of automated decision making in crash scenarios: In a liberal democracy, that has an overlapping consensus with a substantial deontological content it might be unacceptable to sacrifice a fellow road user in order to save a greater number of other fellow road users. However, this might be required in a society with an overlapping consensus with substantial consequentialist content. Now it seems reasonable that a liberal democracy has to ensure that only cars will be importable that respect the local regulation. However, it is less clear if a liberal democracy should allow its ADT producers to export ADT that is respecting the regulations of the importing country but thereby violates core values of its own society.

Other ethical questions can be raised regarding the import and export of ADT between liberal democracies and totalitarian societies. With regard to the import of ADT from totalitarian societies one has to weigh the benefits of free international trade with possible privacy threats and security threats, in particular the security of critical infrastructure in case of political tensions. With regard to the export of ADT there exist the same problems as with the export to other liberal democracies with other core values and additionally one has to weigh the risk of ADT becoming a means of totalitarian suppression.

References

Brändle, C., & Schmidt, M. W. (2021). Autonomous Driving and Public Reason: A Rawlsian Approach. Philosophy & Technology. https://doi.org/10.1007/s13347-021-00468-1
Himmelreich, J. (2018). Never Mind the Trolley: The Ethics of Autonomous Vehicles in Mundane Situations. Ethical Theory and Moral Practice, 21(3), 669–684. https://doi.org/10.1007/s10677-018-9896-4
Rodríguez-Alcázar, J., Bermejo-Luque, L., & Molina-Pérez, A. (2020). Do Automated Vehicles Face Moral Dilemmas? A Plea for a Political Approach. Philosophy & Technology. https://doi.org/10.1007/s13347-020-00432-5

Session 2A: The risk ethics of autonomous vehicles: A continuous trolley problem in regular road traffic

Sebastian Krügel & Matthias Uhl

Deterministic dilemmas resembling the trolley problem still predominate ethical discussions on autonomous vehicles (AVs). This neglects the fact that road traffic is not deterministic, but risky. Connected AVs may be a chance for managing traffic risk more deliberately than impulse-driven manual traffic allows. By focusing on unavoidable accidents, however, the recent ethical perspective sets accident probability to one and solely discusses accident severity. Conversely, the engineering perspective is essentially guided by accident avoidance, and therefore it questions the relevance of the ethics of unavoidable accidents. The focus of engineering is on minimizing accident probability. Both perspectives neglect the fact that any maneuver in regular road traffic constitutes a redistribution of risks that is a function of both accident probability and accident severity. This distribution of risks raises ethically relevant questions that cannot be evaded by simple heuristics such as "hitting the brakes."

To contribute to this ethical discourse, we developed a graphical interface depicting a common traffic situation in a future with AVs operating in mixed traffic. Using this interface, we elicited laypeople's intuitions about the distribution of risks in 29 different traffic situations. In all situations, an AV was depicted between two other road users and the participants could gradually adjust the AV's driving position between the other road users in 99 increments. The middle driving position of the AV between the two other road users minimized the overall accident probability. This probability grew exponentially with deviation from the middle driving position. The shorter the distance to one of the two other road users, the greater the increase in the probability of a collision with this road user.

Using a representative sample in Germany with 1,807 participants, we found that laypeople's preferences significantly differed from mere collision avoidance. On average, the participants positioned the AV always closer to the vehicle with fewer passengers. Surprisingly, the results were very similar when participants had to imagine that they themselves were passengers of the AV. Even then they positioned the AV always closer to the vehicle with fewer passengers. Thus, they expressed their acceptance of a higher accident probability for themselves if this decreases the probability of a more severe accident for others. This is noteworthy with respect to the social dilemma of AVs identified in the context of unavoidable accidents. Based on our sample, it seems that people might be more altruistic in the risky than in the deterministic domain. This divergence in stated preferences provides another reason to shift the ethics of AVs from the

Ryan Wittingslow

In his seminal How to Do Things With Words (1962), J. L. Austin claims that utterances are constituted by three components: locutionary content (what an utterance literally means), illocutionary force (what the utterance conventionally achieves by virtue of those conventions governing the power of the utterance), and perlocutionary outcomes (what the utterance actually achieves). Although Austin's primary concern was linguistic utterances, later theorists have applied Austin's 'speech act theory' to other domains — including, notably, to artworks.

The virtue of applying a speech act approach to artworks is that Austin's tripartite model provides philosophers with a way of distinguishing between what an artwork literally means (the locution) and the force of that meaning (the illocution). However, developments in this domain have, thus far and for the most part, been narrow. Prevailing accounts focus primarily upon artworks or similar expressions with offensive or explicit locutionary content and/or illocutionary force (pornography, certain examples of public statuary, and so on; cf. Langton 1993, Scarre 2o19, Friedell and Liao 2022, etc.). Nonetheless this scholarship shows, in my view, great promise for application in other philosophical domains: not just philosophy of art, but also philosophy of technology, philosophy of engineering, and philosophy of design.

In light of this conviction, in this paper I will apply post-Austinian speech act theory to these three overlapping philosophical domains. Drawing upon Austin's tripartite distinction between locutionary content, illocutionary force, and perlocutionary outcomes, I will furnish a novel and plausible account of technical functions, the ways in which the individual or conventional uses of technical artefacts can deviate from those functions, and the cultural and material conditions that circumscribe the relationship between function and use. Furthermore, I will also highlight similarities and differences between this and competing accounts of technical function, in particular the action theory accounts of Houkes, Pols, Vermaas, et al. (cf. Houkes et al. 2002, Vermaas and Houkes 2003, Pols 2015, etc.).

References

Austin, John Langshaw. 1962. How to Do Things with Words. Oxford University Press.
Friedell, David, and Shen-yi Liao. 2022. 'How Statues Speak'. The Journal of Aesthetics and Art Criticism.
Houkes, Wybo, Pieter E Vermaas, Kees Dorst, and Marc J de Vries. 2002. 'Design and Use as Plans: An Action-Theoretical Account'. Design Studies, 23 (3): 303–20.
Langton, Rae. 1993. 'Speech Acts and Unspeakable Acts'. Philosophy & Public Affairs 22 (4): 293–330.
Pols, Auke. 2015. 'Affordances and Use Plans: An Analysis of Two Alternatives to Function-Based Design'. AI EDAM 29 (3): 239–47.
Scarre, Geoffrey. 2019. 'How Memorials Speak to Us'. In Philosophical Perspectives on Ruins, Monuments, and Memorials, 21–33. Routledge.
Vermaas, Pieter E., and Wybo Houkes. 2003. 'Ascribing Functions to Technical Artefacts: A Challenge to Etiological Accounts of Functions'. The British Journal for the Philosophy of Science 54 (2): 261–89.

Session 2B: A Taxonomy For Artifact Function Pluralism

Clint Hurshman

This paper aims to motivate and structure a pluralist view of artifact functions.

Theories of artifact functions diverge in the role that they assign to designers' intentions (see Preston 2009) and, relatedly, in the kind of property that they take functions to be, viz. whether they depend on ascription or not (van Eck and Weber 2014). Unsurprisingly, they have also used the notion of function to perform a variety of kinds of discursive work, including: describing the phenomenology and social epistemology of artifact use; justifying the use of artifacts for given purposes; explaining the etiology of the forms and uses of artifacts; and explaining malfunction. There is little reason to think that a single account of function in this context will be adequate to perform such varied tasks, giving us some reason to follow Perlman (2009) in accepting a "pragmatic teleo-pluralism" about artifact functions. However, the structure of such a pluralistic view has not been explicitly elaborated: what will the parts of such a pluralism be—and what "functions" will they serve?

This paper aims to contribute to the literature on artifact functions by pulling apart these discursive tasks and comparing how existing theories address them. It takes an ameliorative approach, asking: how does the work that the notion is being used to perform inform the function that should be ascribed to the artifact? Focusing on the tasks of describing the social epistemology of artifact use and explaining the etiology of artifacts, I examine the implications of different accounts for these discursive tasks, depending on whether they are intentionalist or nonintentionalist and whether they treat functions as depending on ascription.

Finally, I argue that another task warrants further attention, namely informing intervention. The way that the function of an artifact is delineated entails prescriptions about how to intervene when artifacts have undesired effects. This "interventionist task" of functions has been acknowledged in the context of biological functions (e.g. Garson 2010), and less explicitly in the context of artifacts by Preston (2009) and Van Eck and Weber (2014). I argue that performing this task adequately requires a non-intentionalist, function-as-property account.
By mapping out a pluralistic picture, this paper aims to facilitate further conversation between accounts often seen as competing.

References

Buller, D. J. (1997). Individualism and evolutionary psychology (or: In defense of "narrow" functions). Philosophy of Science, 64(1), 74-95.
Houkes, W., & Vermaas, P. (2010). Technical Functions: On the use and design of artefacts (Vol. 1). Springer.
Perlman, M. (2009). Changing the mission of theories of teleology: DOs and DON'Ts for thinking about function. In U. Krohs & P. Kroes (Eds.), Functions in biological and artificial worlds: Comparative philosophical perspectives (pp. 17-36). MIT press.
Preston, B. (2009). Philosophical theories of artifact function. In A. Meijers (Ed.), Philosophy of technology and engineering sciences (pp. 213-233). North-Holland.
Van Eck, D., & Weber, E. (2014). Function ascription and explanation: Elaborating an explanatory utility desideratum for ascriptions of technical functions. Erkenntnis, 79(6), 1367-1389.

Session 2C: Animal Models and Human Diseases: The Ethics of Genome Editing Animals to Cure Humans

Mattia Pozzebon

"An animal model is a non-human species used in biomedical research because it can mimic aspects of a biological process or disease found in humans" [1]. By using animal models, it is possible to understand the development in humans of rare diseases such as the Timothy syndrome, the Stargardt disease, ALS, etc. It allows significant advances in medical research that would otherwise be impossible. Indeed, the same research carried out on human beings would be limited due to ethical prohibitions [2]. Among animal species, pigs are one of the most suitable since they are genetically and physiologically the closest to humans after primates [3,4].

The use of animal models "in the biomedical field has been consolidating in recent years thanks to the development of somatic cell nuclear transfer (cloning) and genome editing technologies (Zn Fingers, TALENS and CRISPR/Cas9)" [5]. By genetically engineering the cells used for cloning, transgenic animals are generated. These animals carry the specific disease that will later be studied. Therefore, somatic cloning technique (i.e., somatic cell nuclear transfer), combined with genome editing technologies, results in diseased animals beneficial for developing treatments for humans.

The aim of this presentation is then to discuss the welfare implications for the animals involved. While acknowledging the undeniable value of animal models for biomedical research [6], why should it be ethically permissible to breed disease-carrying animals to seek treatments for humans? A fortiori, by begetting diseased animals that would otherwise never have been born. Given the non-identity problem, if we recognise that the existence of these animals is impaired so as to be not worth living, should we deem genome editing impermissible? However, at the same time, it is also important to answer a further question: Is human suffering more valuable than animal suffering? By answering affirmatively, should we then admit the admissibility of animal models? Would it then be ethical to discuss the further genetic modification of these animals in order to reduce experienced suffering, e.g., by genetically disenhancing them [7]?

References

1. Ostrander EA (2023). Animal Model. National Human Genome Research Institute, January 9th. https://www.genome.gov/genetics-glossary/Animal-Model#:~:text=Definition&text=An%20animal%20model%20is%20a,or%20disease%20found%20in%20humans.
2. Mukherjee P, Roy S, Ghosh D, et al. (2022). Role of animal models in biomedical research: a review. Lab Anim Res 38:18. https://doi.org/10.1186/s42826-022-00128-1
3. Schook LB, Collares TV, Darfour-Oduro KA, et al. (2015). Unraveling the Swine Genome: Implications for Human Health. Annual Review of Animal Biosciences 3(1):219-244. https://doi.org/10.1146/annurev-animal-022114-110815
4. Crociara P, Novella Chieppa M, Vallino Costassa E, et al. (2018). Motor neuron degeneration, severe myopathy and TDP-43 increase in a transgenic pig model of SOD1-linked familiar ALS. Neurobiology of Disease 124:263-275. https://doi.org/10.1016/j.nbd.2018.11.021
5. Avantea. Animal Models. https://www.avantea.it/en/animal-models/
6. Robinson NB, Krieger K, Khan FM, et al. (2019). The current state of animal models in research: A review. International Journal of Surgery 72:9-13. https://doi.org/10.1016/j.ijsu.2019.10.015
7. Devolder K, Eggel M (2019). No Pain, No Gain? In Defence of Genetically Disenhancing (Most) Research Animals. Animals 9(4). https://doi.org/10.3390/ani9040154

Session 2C: Techno-fixing hemophilia away: an ethical assessment of arguments for gene therapy

Lieke Baas, Rieke van der Graaf, Karina Meijer & Annelien Bredenoord, for the SYMPHONY consortium

Since the beginning of gene therapy development, hemophilia has been considered as the perfect test case for validating general gene therapy principles. Similarly, the hemophilia community has embraced gene therapy's great promises. In addition to the promise of providing a definitive cure, several arguments are provided in favor of the development of gene therapy for hemophilia. These do not only concern medical arguments, but also ethical arguments, such as the hope that gene therapy will solve the global inequality in access to treatment and that it will provide people with hemophilia with more freedom.

In this paper, we analyze the arguments put forward in the debate on hemophilia gene therapy. We start with inventorying the arguments used in favor of gene therapy in relevant documents, such as policy documents and clinical trial protocols. We argue that some of these arguments, such as the potential that gene therapy will solve the global gap in access to treatment, are in fact a technological fix, that is, using a technological innovation to solve problems that are inherently social or political, instead of technical. The current treatment gap is an example of this, as it is does not result from a lack of treatment options but from the unequal distribution of wealth. It has been argued that technological fixes can have several ethical consequences, such as letting commercial interests prevail over patients' best interests, and their narrow scope can be in the way of finding potentially superior solutions to the problem. Drawing from the sociology of expectations, we evaluate the ethical consequences of technological fix narratives for gene therapy.

Session 2D: Bringing back words from their metaphysical to their everyday use: How Wittgenstein and cognitive-functional linguistics can help us to better understand and teach engineering communicative practice

Rick Evans

At a prior fPET conference, I argued that the origins of modern engineering communicative practice could be found in Wittgenstein's idea of the "language-game." According to Wittgenstein (1953), a "language-game is meant to bring into prominence the fact that the speaking of a language is part of an activity or a form of life" (PI §23). My claim was that engineers develop a specialized and complex variety of language, a language-game (many language games actually), that is not only associated with engineering practice, but also is necessary to practice as an engineer. Because Wittgenstein practiced what he preached, "don't think, but look" (PI §66), he saw that there were different "forms of human life," e.g., life as an engineer, and different ways of using language, language-games emergent from those forms.

In a 1944 letter to Norman Malcolm, Wittgenstein states "… what is the use of studying philosophy if all it does for you is to enable you to talk with some plausibility about some abstract questions (Malcolm 1984). Following his lead, I might ask of myself -- "What is the use of studying Wittgenstein, if all that it enables me to do is talk 'with some plausibility' about the origins of modern engineering communicative practice?" The later Wittgenstein did not embrace new theory-making. Instead, Wittgenstein suggests that "What we should do is to bring words back from their metaphysical to the everyday" (PI §116).

As a sociolinguist, I advocate focusing on everyday language-in-use. Like Wittgenstein, I also say "don't think, but look." But, when we look at engineering communicative practice, what do we see? We see language games that are highly specialized in relation to various and different disciplinary practices – "the meaning of a word is in its use" (PI §43). We see that the only way to learn these language games is to engage in those practices – "adults talk to him and teach exclamations" (PI §244). And we see a constant struggle between convention, "every course of action can … accord with the rule" and change, "no course of action could be determined by a rule" (PI §201). In other words, we see language-in-use as situated action, we see that learning can only happen through engaging in that situated action, and we see that language-in-use evolves through obeying and breaking rules. Important insights concerning language-in- use generally and engineering communicative practice.

Only recently through the discoveries of cognitive-functional linguistics are we able to begin to appreciate these insights that Wittgenstein offered to us. My paper/presentation begins with an articulation of those insights. I support each of the insights with evidence generated by cognitive-functional linguistics. Then, because I have for over 20 years been engaged teaching engineering communicative practice, I describe how these insights can help us to better understand how we might teach engineering communicative practice or those certain language games. Such an understanding should help us to design a more useful pedagogy. I will briefly offer what I believe are the key elements of that pedagogy.

Session 2D: Epistemological challenges for multidisciplinary design

Martin Stacey & Claudia Eckert

Multidisciplinary design involves communication across disciplinary boundaries that presents both practical and philosophical challenges. Engineering designers and other participants in engineering design projects need to talk to each other, and they need to collaborate in large-scale development processes where they need to work together and separately, generating and supplying information and participating in decision-making. Multidisciplinary collaboration requires mutual respect that is sometimes lacking, but also sufficient mutual understanding for effective cooperation (see Eckert, 2001; Eckert and Stacey, 2021).

Failures in multidisciplinary collaborations can come from epistemic problems: failures not just to understand the content of other participants' knowledge, but failures to understand the epistemic relationships between models and what they represent, and the epistemic relationships between what other participants know and what they are trying to decide or establish in their design work. In our empirical research on design processes we have observed some of these failures; we have observed the participants in the design processes misinterpret problems stemming from these failures as colleagues being awkward, or as communication issues. Engineers dealing with multidisciplinary projects in large engineering companies have discussed similar issues with us.

We argue here that while problems stemming from lack of mutual understanding go well beyond epistemological issues, looking at them from an epistemological perspective is not merely intellectually interesting but gives us an approach to enabling better collaboration in multidisciplinary teams.

This is becoming both more complex and more urgent. While engineering products have combined mechanical systems with electronics and software for the last 50 years, the way mechanical engineers, software developers and electrical engineers approach design is still very different and their activities in organisations are not well integrated. One of the big challenges over the next decades is to integrate different engineering disciplines with different traditions into coherent and effective development processes. Nor is software development epistemologically homogeneous. Mechatronic systems can embed software with very different epistemic foundations, including provable algorithms, statistical inference, heuristic rules or assertions in symbolic AI systems, and decision procedures induced from data in neural network systems.

Epistemological mismatches present challenges for developing methods for multidisciplinary design that integrate the work of different disciplines. Many companies are still dominated by the original discipline of the product, e.g. car companies are largely managed by mechanical engineers, and construction projects by civil engineers. With this come a dominance of the tools and methods of these domains; and an unbalanced power relationship between what is still seen as the “home discipline” and the new disciplines.

References

Eckert, C.M. (2001). The communication bottleneck in knitwear design: analysis and computing solutions. Computer Supported Cooperative Work, 10, 29-74.
Eckert, C.M. and Stacey, M.K. (2021). Overconstrained and Underconstrained Creativity: Changing the Rhetoric to Negotiate the Boundaries of Design. In L.T.M. Blessing, A.J. Qureshi and K. Gericke (eds.) The Future of Transdisciplinary Design. Springer, pp 207-219.

Session 2E: The Rise and Fall of the Code of Ethics of the Chilean Engineering Council: History and Vicissitudes

Patricio Quintana Gallo & Carlos Verdugo Serna

The Code of Ethics of the Chilean Engineering Council has been the subject of examination and discussion in some international and national publications starting in 2001. An example of these are the books “Ethics in the Engineering Profession: Engineering and Citizenship” written by Karl Mitcham and Marcos García de la Huerta, and “Ethics and Science: Cultural Frontiers”, edited by Jesús Valero Matas and others, including Karl Mitcham. In these contributions, however, the analysis of the content and the principles that appear to sustain such a code is not extensive. Considering the latter, the main objective of this work is to conduct a deeper study about the relationship between ethics and engineering in Chile, by means of a critical examination of the specific contents of the Code of Ethics of the Chilean Engineering Council and its fundamentals. A second objective is to foment the dialogue and the more sustained interdisciplinary study between engineers and philosophers of the Chilean universities. The work will attempt to identify and expose the values and ethical principles that could be explicit or not in the aforementioned Code of Ethics, such as the classical principles of non-maleficence, beneficence, autonomy, and justice, as well as those related to technical-scientific systems, such as truthfulness, privacy, consent, responsibility, confidentiality, and loyalty. In this presentation, we hope to also discuss with engineers and engineering students the knowledge that they could have about this Code and the real importance that they attribute to the education of engineers and their professional performance.

Session 2E: A comparative institutional analysis of engineering ethics in the Netherlands and the United States

Ali Dizani, Behnam Taebi, Ibo van de Poel & Amineh Ghorbani

The first discussions of engineering ethics date back to half a century ago. This paper focuses on the development of engineering ethics in the national context. In order to design strategies for the development of engineering ethics, national policymakers must be able to recognise contextual factors and their impacts.

Several countries have played a crucial role in the development of engineering ethics, including the Netherlands and the United States. In this paper, we will review the development of engineering ethics in these pioneering countries. The research aims to compare the institutional context of engineering ethics in these countries, with the aim to identify the key institutions that influence its development at the national level. These could help to offer recommendations for national policy-making in other countries with an interest in engineering ethics.
The topics analysed in these two countries include the underlying reasons for the first development, the actions taken by actors in the development process, the way this process was carried out, and the impact of contextual factors, specifically institutions, on this process. Data was collected through a literature review and structured interviews. To analyse the activities carried out in the development process, an earlier framework developed for National Engineering Ethics (NEE) (proposed by Dizani, Under Review) was used. This framework builds on the work of Elinor Ostrom.

The comparative analysis highlights many dissimilarities in the social contexts of these countries with regard to the development of engineering ethics. For instance, the beginning of engineering ethics debates in the United States started with scandals and engineering accidents, while in the Netherlands it was the technological progresses that provided a suitable social context for the engineering ethics movement. The difference between the nature of engineering professions is another example. The engineering professional organisations in the USA have more autonomy and authority, compared to the Netherlands. Additionally, it was clear that governments, academia, and professional organisations played distinct roles in the development process. In the Netherlands, engineering ethics has been strongly institutionalised in the first years of its development. For instance, there were various government advisory bodies, especially the Rathenau Institute that aimed at assessing the effects of technology on life. In the United States, on the other hand, government played a less prominent role but engineering professional associations such as IEEE or ASCE were more crucial; their involvement was, for instance, in drafting the codes of conduct for professions or asking universities to include ethics courses in engineering education.

Session 2F: Knowledge-policy interfaces for wicked problems. The Co-production of Digital Sequence Information governance at the Convention for Biological Diversity

Adam McCarthy

Governing rapidly evolving technologies to fulfil the needs and values of diverse stakeholders is non-trivial, especially when such technologies relate to urgent, wicked problems such as biodiversity loss and climate change. Nonetheless, theoretical advances such as participatory technology assessment and Inclusive Innovation aim to integrate diverse values into technology governance and link technological transformation to policy at various scales (Mission-oriented innovation policy, SDGs etc.).

An example of this is current debates about digitalisation in biosciences at the UN Convention for Biological Diversity (CBD) where integrating evidence and diverse values to achieve desirable governance outcomes is challenging, since the subject is technical, and decisions have significant consequences. It is widely recognised that genome sequencing-enabled digitalisation undermines the CBD norms of ‘fair and equitable sharing of benefits arising out of the utilisation of genetic resources’ because the definition of genetic resources was limited to physical samples in 1992. In other words, technological change in the biosciences escaped a governance mechanism to link biotechnology innovation to conservation in ‘megadiverse’ countries and parties to the CBD are currently deciding whether, and how to re-establish biological Digital Sequence Information (DSI) under its governance regime. The governance of DSI involves managing the interests of a diverse group of stakeholders, encompassing both North-South and public-private dimensions. The CBD process also emphasizes the role of Indigenous peoples as Traditional Knowledge holders and as potential recipients of benefits from DSI governance.

These interests feed into the CBD Science-Policy Interface where different institutional norms and power dynamics interact to understand, and subsequently decide upon, the issue of DSI governance. This paper utilises Jasanoff’s notion of Co-production to explore the features of the DSI Science-Policy Interface, exploring to what extent a formal technology governance process can be inclusive of different needs, values, and rights in the context of global, wicked problems. To achieve this, I use a critical review of Co-production literature to develop a theoretical foundation for further analysis. Following this, documentary analysis is used to conceptualise the CBD’s approach to DSI according to Co-production and the contribution of academia is analysed using bibliometric analysis. These insights are supplemented by the results of participant observation at the recent COP15. Because the CBD is a global convention, it operates under UN rules which are ostensibly inclusive and aim to ensure representative and pluralistic participation. This paper explores this participation in practice while assessing the value of Co-production as an underlying theoretical framework.

Preliminary findings suggest that the notion of a Science-Policy Interface may be limited considering the diverse knowledge and values interacting in this case. Therefore, this paper advances the concept of a Knowledge-Policy Interface, exploring its relevance and dynamics in the context of this unfolding issue. Investigating this Knowledge-Policy Interface should provide useful insights for the study of participatory technology governance as well as for the actors involved in the case. The findings also form the foundation for further empirical investigation, exploring whether values of inclusion, ethics, and responsibility can be integrated into international policymaking for wicked problems.

Session 2F: How to develop principles for fair access and benefits of Digital Sequence Information for scientists and engineers

Bob Kreiken & Lotte Asveld

Data-centric biology is taking a great leap, thanks to developments in omics technologies and the ever expanding genomic sequencing databases which consist of so-called Digital Sequence Information (DSI). DSI is derived from the sequencing of living organisms such as microbes, plants and animals. These data in turn feed into machine learning approaches, which help to sophisticate engineering of living organisms. These developments go hand in hand with the Open Science movement that urges scientists to share their data as much as possible, following the principle that more data is always better and will lead to more scientific insights.

However, these developments risk increasing the digital divide between scientists with access to sufficient resources to generate and analyse DSI in some areas of the Global North, as compared to scientists in resource-limited regions, mostly in the Global South (Leonelli, 2016). Without enabling conditions 'open access to data' does not guarantee 'fair access to data'. The inability to ''self-use'' DSI (Scholz et al., 2021) means that researchers in the Global South have less scientific agency over DSI and priority-setting in scientific research.

This disparity between the ability to access and create benefits from data has led countries in the Global South to demand compensation for the use of DSI that originated from their region. Although many agree that it is fair to compensate and emancipate resource-limited but biodiversity rich countries (Bagley, 2016), the main issue is how to arrange this. At the Conference of the Parties (COP15) of the Convention on Biological Diversity (CBD) past December, Parties agreed to develop a multilateral benefit-sharing fund to which users of DSI would contribute. This fund would then be used to invest in biodiversity conservation and scientific capacity building in the Global South. However the question remains how to distribute the yield of such a multilateral fund?

We argue that a fair distribution should at least take into account the different capabilities of scientists in different countries in generating and analysing data. In this paper we will focus mainly on academic scientists and engineers in various regions, leaving other actors such as indigenous peoples and companies out of the distribution picture for now. In developing principles for a fair distribution of benefits from DSI, we propose these three conditions:

• The burden of execution shouldn't be put solely on individual researchers and research organizations in the Global North, but they have a duty to contribute
• Fair distribution of benefits should contribute to fair distribution of scientific agency
• Fair distribution of access implies that management of data is restructured to reflect the needs of the Global South

Session 3A: Ecosystem-education and universities' social responsibility

Gunter Bombaerts, Diana A. Martin, Neelke Doorn, Zoë Robaey, Jeremy Mantingh & Suleman Audu

Panel focus

Technical universities increasingly organize education in cooperation with their ecosystems. This "ecosystem-education" goes by many names, such as community-based, project-based, design-based, challenge-based, mission-based, and so forth. All these approaches have in common that they focus less on pre-given knowledge transfer from teacher to passively receiving student. They focus more on activating students to solve a complex and open-ended challenge of a real-life stakeholder, often in the network (or eco-system) of the university. See video1 and video2 for some examples.

Of course, this education is not neutral. Students have interactions with external stakeholders (companies, communities, ngo's, ...). In these interactions, all kinds of value and norm discussions take place. Most of the times, these discussions are not very system-critical and are an "interesting" deepening of the participants' reflection of their moral role as students, teachers, engineers, public servants or citizens. Sometimes however, discussions are critical towards the ongoing activities or within the existing eco-system. Discussions then initiate potentially sensitive reflection processes on existing norms of the university and the eco-system partners. Companies that collaborate in this ecosystem-education might receive very fundamental critiques by students. But also the organizing university can be forced to reflect on its social responsibility.

This raises several questions of the ecosystem-education. We want to focus on the role of this education on the universities' social responsibility. Should this ecosystem-education avoided ethically sensitive situations; or should it, on the contrary, allow or even deliberately stimulate these critical situations?

This session will analyze this question. It will consider the contextualities in situations that could be used to argue for a specific approach of avoiding, allowing or stimulating). What are the relevant culture and norms at universities? How does a particular university frame itself, in light of social responsibility, both in its mission and in how it adopts concrete societal positions? What are the visions and aims of the external organization (think of Google, Shell, University Rebellion, ...) and how should ecosystem-education cope with these? And so forth.

Panel participants
- Gunter Bombaerts and Diana Martin (TU Eindhoven, the Netherlands)
- Suleman Joseph Audu (Engineers Without Borders)
- Neelke Doorn (TU Delft, the Netherlands)
- Non-European participant (tbd)

Schedule
• Introductory presentation (Gunter Bombaerts and Diana Martin)
• Panelists give short reaction/statement to the presentation
• Discussion with the public

Session 3B: Why We Need a Phenomenology of Brain Computer Interface-mediated Communication

Bouke van Balen

In this paper, I argue for the need of an interdisciplinary and empirically informed phenomenology of communicating through Brain Computer Interfaces (BCIs). BCIs are becoming more used as a form of Augmentative and Assistive Communication technologies (AAC-tech), to provide people without a natural speaking voice with a manner of communication. Perhaps the most promising application of BCI as AAC-technology are implantable speech-BCIs, devices that record and translate brain activity of attempted speech into words on an external interface. This technology could enable individuals who have lost control over their muscles, such as ALS-induced Locked-In patients, to regain their communicative abilities.

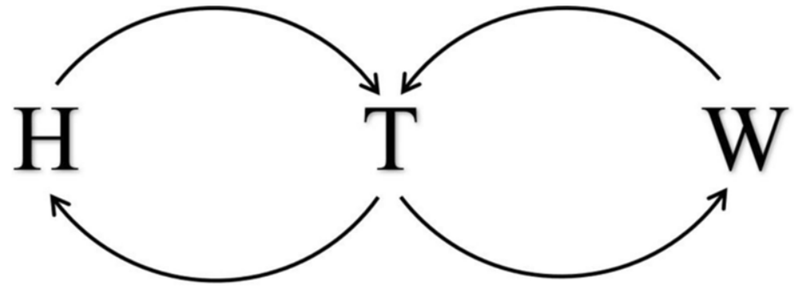
In the most recent breakthrough in the development of implantable speech-BCIs it was claimed that BCIs have the potential to 'restore' communication (Chang et al., 2022). However, I argue from a postphenomenological and enactive 4E-perspective that this is not the full picture, and that we should be wary of these kind of restoration-claims. The idea that BCIs can restore communication rests on two invalid instrumental assumptions. One is that BCI is a neutral tool, the other is that communication is an instrumental act, by which I mean a functional exchange of information. I argue that speech-BCIs are not neutral tools but mediate the experience of communication, thereby installing a new kind of communication. Zooming in on this mediated experience of communication by drawing on personal testimonies from AAC-users, I show, in line with recent work from Van Grunsven & Roeser (2022) that the affordances of this mediated form of communication are instrumental, leading to a limited experience of communication that misses out on the enactive, embodied and social process that interpersonal communication is.

After establishing that the experience of communicating through an implantable speech-BCI differs from communicating through conventional speech, the question rises how this experience is different. Surprisingly, the phenomenology of BCI-mediated communication, and phenomenologies of neurotechnologies in general, have received little attention in the ethical debate around BCIs. I argue that this phenomenology is important to understand for three reasons. One, for scientific reasons, it can help to increase the understanding of the relationship between neurotechnology, the brain, embodiment, and our subjective experience. Two, for moral reasons, the experience of the users must be included as a matter of epistemic justice and as a matter of epistemic access to moral knowledge. Three, future research and development in the field of BCI, neurotechnology, and AAC-technology needs to be informed by the scientific and moral knowledge that reason one and two provide. Concluding, neuroscientists, engineers, designers, and philosophers share the interdisciplinary task to pursue an empirically informed phenomenology of BCI-mediated communication.

Session 3B: Denominating the Movements in the Lemniscate of Technology

Lars Botin & Tom Børsen

The hermeneutic lemniscate (of technology) was originally introduced by Olya Kudina (2021) and inspired by Hans Georg Gadamer's hermeneutic circle, Don Ihde's material hermeneutics and Peter-Paul Verbeek's human-technology-world relations. In this paper we will try to qualify the four movements in the lemniscate to further the understanding of how technology is interdeterment for how the other components in the lemniscate, humans and world, are under constant pressure and change.



The first movement (1) is from human to technology and is concerned with how we design and manufacture technology. We have identified this movement as tinkering. Users, designers, engineers etc. tinker with technologies in multistable ways to achieve something. The second movement (2) regards the intentional and unintentional implications of technology in the world. Technologies give access to the world. The third movement (3) is the responsiveness of the world in relation to technological innovation of sustainable and responsible solutions. The world produces responses, and it makes queries, proposals and demands to technologies. The fourth movement (4) regards the potentials and capacities of technology to form humans, and their interpretations of the world and themselves. Technologies shape our experiences, behaviors, understandings, and bodies.
Humans tinker and are shaped by and with technology (1 & 4). It is what Gilles Deleuze has defined as a constant becoming, characterized by effects and events, and framed by processes and verbs. World, on the other hand, is substantial being of causes (2 &3), framed by things and adjectives/substantives (Deleuze 1990/2004). In this way the lemniscate can be divided into two distinct spheres, even though interdependent and intertwined.

On the left-hand side effects and events are driven by tinkering and shaping, and on the right-hand side causes manifest in tangible things. Deleuze writes: "On one hand there are singular proper names, substantives, and general adjectives which indicate limits, pauses, rests, and presences (2 & 3); on the other there are verbs carrying of with them becoming and its train of reversible events and infinitely dividing their present into past and future (1 & 4)" (Deleuze 1990/2004, p. 30).

Kudina stresses that there is no chronology in the lemniscate, and you can enter wherever is appropriate, furthermore that technology is not central, but aligned with humans and world. We see technology as gateway and -keeper in between the two spheres of effects and causes, hence determent for how humans and world co-constitute in past, present and future by and with technology.

Keywords: Lemniscate of technology, Becoming and Being, postphenomenology, Deleuze

References:
Deleuze, G. (1990/2004): The Logic of Sense. London/New York: Continuum
Gadamer, H.-G. (1960/2011): Truth and Method. London/New York: Continuum
Ihde, D. (1998): Expanding Hermeneutics. Visualism in Science. Evanston, Illinois; Northwestern University Press
Kudina, Olya (2021): "Alexa, who am I?": Voice Assistants and Hermeneutic Lemniscate as the Technologically Mediated Sense-Making. Hum Stud, 44, 233-253
Verbeek, P.-P. (2011): Moralizing Technology. Understanding and Designing the Morality of Things. Chicago/London: The Chicago University Press

Session 3C: Complex Fairness

Bauke Wielinga

Many recent papers take a statistical approach to algorithmic fairness, which aims to ensure fair treatment between two groups based on the rates at which an algorithm (accurately) categorizes each group Carey & Wu, 2022), e.g. to the category of good candidates, or potential fraudsters. This paper will deal with three problems that this "statistical fairness" approach to algorithmic fairness is argued to face due its abstract nature (Selbst et al., 2019): Firstly, it presupposes a framing of the situation which may obscure fairness-relevant contextual considerations. Secondly, this lack of contextual information may wrongly lead to the conclusion that an algorithm can be exported from one context to another, while there are important contextual differences. Finally, the statistical fairness approach implies an abstract mathematical conception of fairness, which ignores aspects of fairness which are difficult to quantify, such as procedural fairness.

This paper will work towards addressing these problems by construing fairness through the lens of Michael Walzer's notion of "complex equality" (Walzer, 2010). According to Walzer, different social goods, such as money, healthcare and political power, are each associated with their own social sphere, and their own criteria for how they ought to be distributed. Insofar as these criteria ask that each good not be distributed based on how much one has of another good, justice then requires the autonomy of different spheres of justice, in such a way that "no citizen's standing in one sphere or with regard to one social good can be undercut by his standing in some other sphere, with regard to some other good" (Walzer, p. 19).

To illustrate how complex equality can be applied to the debate on algorithmic fairness, and to show how it helps avoid the above problems, two examples from the domain of social security and work will be considered: on the one hand, a hypothetical social security fraud detection system, and, on the other, a CV-scanning algorithm. I will argue that applying complex equality to algorithmic fairness necessitates determining what good is being distributed, and as holders of which goods the distributors and receivers should be understood. Answering these questions in terms of social goods requires a thorough and critical understanding of the situation at hand, thereby alleviating the abstractness of statistical fairness. In the examples, the nature of the different social goods being distributed (social security and employment, respectively), and the different populations under which they are to be distributed (applicants for social security and for jobs) will entail different requirements of fair distribution.

References

Carey, A. N., & Wu, X. (2022). The statistical fairness field guide: Perspectives from social and formal sciences. AI and Ethics. https://doi.org/10.1007/s43681-022-00183-3
Selbst, A.D., Boyd, D., Friedler, S.A., Venkatasubramanian, S., Vertesi, J. (2019): Fairness and abstraction in sociotechnical systems. In: Proceedings of the Conference on Fairness, Accountability, and Transparency. FAT* '19, pp. 59–68. Association for Computing Machinery, New York (2019).
Walzer M. (2010). Spheres of justice : a Defense of Pluralism and Equality ([Nachdr.]). Basic Books.

Session 3C: Justice and Smart Societies – Algorithmizing Fairness

Florian Richter

The paradigm of classic technology is to control and regulate processes to ensure repeatability. In the so-called transclassic technology we face an "intellectualization of the environment", for example, in ubiquitous computing. (Hubig 2007, 43) Technological strategies and procedures are part of the environment. The implementation of smart technologies into the environment takes place on a large scale in so-called "smart cities" or smart societies ("Society 5.0"). (Hitachi-UTokyo Laboratory 2018, 10/11) Smart technologies should make cities, governments, and societies more efficient, effective, and sustainable and thus, more livable for humans. (Bibri 2018), (Kumar et al. 2021), (Zhong 2021) So far, ethical aspects of smart cities, like e.g., "surveillance", "data privacy", and "sustainability" have been examined. (Ziosi et al. 2022) However, questions of justice have not been sufficiently scrutinized. Whereas, Coeckelbergh discusses equality and distributive justice in his political philosophy of technology, he does not investigate justice with regard to smart cities and societies. (Coeckelbergh 2022, 43-48)

If we assume that through the collection and algorithmic processing of large amounts of data, a more personalized conception of fairness would be possible, then we already commit ourselves to one conception of fairness. It is a form of moral desert that assumes that we deserve X, because we need it to live a good life. And these (latent) needs of the citizens could be predicted by e.g., AI-based systems. (Needs are here taken in a general way by covering any aspect (physical, emotional, social, etc.) that one might need to achieve a certain kind of well-being.) I will argue that the conception of justice in smart cities or societies is not a realistic model for distributive justice; instead, another model is proposed that takes algorithmizing fairness as a starting point that allows the evaluation of different conceptions of fairness for different (smart) contexts. Since also debates of political philosophy are echoed in the formalization of fairness in machine learning (Binns 2018), this topic needs further analysis in smart environments.

Not only that it might be particularly difficult for systems to address the (latent) needs of the citizens "perfectly" or even to detect them. There needs to be first a general debate about which might be the best way to cover the basic needs: should it be a universal basic income, a social welfare state, or should it be based on competition (merit-based) and economic incentives (neoliberalism, forms of libertarianism)? Even a liberal state with a certain social welfare system faces the challenge how "compensating benefits for everyone, and in particular for the least advantaged members of society" (Rawls 1999, 13), can be distributed and how those benefits can be used by its citizens or how those primary goods can be converted "into achievements of well-being" (Sen 1995, 27). Consequently, it is crucial to make certain conceptions of justice available and explicit to evaluate them for different contexts. It is thus proposed that technological systems accompany the discourse by algorithmizing conceptions of fairness. It will make processes more transparent and first and foremost possible to discuss and examine which kind of conception of fairness might be adequate in a certain situation without imposing one. Directions of future research are indicated by outlining how empirical studies can complement the conceptual framework.

References

• Bibri, S.E. (2018). "A foundational framework for smart sustainable city development: Theoretical, disciplinary, and discursive dimensions and their synergies." In Sustainable Cities and Society, 38, 758-794.
• Binns, R. (2018). "Fairness in machine learning: Lessons from political philosophy." In Conference on Fairness, Accountability and Transparency, 149-159).
• Coeckelbergh, M. (2022). The Political Philosophy of AI: An Introduction. Cambridge.
• Hitachi-UTokyo Laboratory. (2018). Society 5.0: A People-centric Super-smart Society, SpringerOpen.
• Hubig, Christoph. (2007). Die Kunst des Möglichen II. Ethik der Technik als provisorische Moral. Bielefeld.
• Kumar, A., Sharma, S., Goyal, N., Singh, A., Cheng, X., Singh, P. (2021). "Secure and energy-efficient smart building architecture with emerging technology IoT." In Computer Communications, 176, 207-217.
• Oberascher, M., Rauch, W., Sitzenfrei, R. (2022). "Towards a smart water city: A comprehensive review of applications, data requirements, and communication technologies for integrated management." In Sustainable Cities and Society, 76.
• Rawls, John. (1999). A Theory of Justice, revised edition, Harvard University Press, Cambridge (Mass.).
• Sen, Amartya. (1995). Inequality Reexamined, HUP Cambridge (Mass.).
• Zhong, C., Ng, S.T., Skitmore, M. (2021). "An interdependent infrastructure asset management framework for high-density cities." Proceedings of the Institution of Civil Engineers: Municipal Engineer, 174 (3), 180-190.
• Ziosi, M., Hewitt, B., Juneja, P., Taddeo, M., Floridi L. (2022). "Smart cities: reviewing the debate about their ethical implications." In AI & Soc (2022)

Session 3C: Differences in perceptions of justice between police and AI-based automated enforcement

Verity Truelove, Oscar Oviedo-Trespalacios, Kayla Stefanidis & Levi Anderson

Numerous jurisdictions worldwide are considering implementing AI-based automated enforcement cameras that can detect drivers illegally using their phone while driving. While these cameras have the potential to increase drivers' chances of being caught and punished for violating this rule, it is important to consider how the technology may impact perceptions of justice ort he enforcement practice and ultimately, engagement in phone use while driving. The application of procedural justice theory can be used to understand perceptions of justice ort he technology. Based on procedural justice theory, it is suggested that if people are perceived ort treated fairly by the enforcement initiative, they will have more respect towards the law, leading to higher compliance (Tyler, 1990). To understand the procedural justice impact of the mobile phone detection cameras, it is important to consider this impact against police enforcement practices, which had been the primary enforcement practice ort he road rule violation prior ort h implementation of the AI-based phone enforcement technology. Therefore, this project aimed to understand the procedural justice impact of the phone detection cameras compared with police enforcement of phone use while driving. Two studies were conducted to address this aim. In Study 1, interviews with 26 police officers from Australia were conducted. The data was analysed using a theory led approach (Howitt & Cramer, 2014), with codes developed based on the four principles within procedural justice theory (fairness, voice, trustworthiness and respect). Overall, the interviews suggested that police enforcement allowed drivers more of an opportunity to give voice ort he experience, while the enforcement cameras provide a more neutral enforcement experience. Next, a quantitative survey was conducted with 486 Australian drivers. From the drivers' perspective, it was found that police enforcement was generally viewed as significantly more fair ($p < .001$), trustworthy ($p < .001$), and respectful ($p < .001$) than the cameras. However, there was no significant difference in perceived voice between police and camera enforcement ($p = .942$), suggesting that drivers believe the process for voicing complaints is similar. It can be suggested that automated cameras do not improve the overall social standing of the road rule enforcement yet can be useful for imparting justice. The results from this study have important implications for the application of mobile phone detection cameras worldwide.

References

Howitt D, Cramer D. (2014). Introduction to Research Methods in Psychology. UK: Pearson Education Ltd
Tyler TR (1990). Why people obey the law. Yale University Press, New Haven

Session 3D: The Harm of Social Media to Public Reason

Paige Benton & Michael W. Schmidt

It is commonly agreed that some features of social media, like so-called echo chambers and filter bubbles, are detrimental to liberal democracies. One reason for this assessment is the claim these features are somehow causally responsible for a process of political polarization, fueling ignorance, hatred and hostility. However, social media has also huge potential in offering vivid channels of communication that overcome distances in space, time and social stratification, and thus might foster mutual understanding and social cohesion.

With reference to John Rawls's political liberalism, we offer an additional explanation of why some features of social media indeed can be seen as being responsible for the violation of necessary norms of liberal democracies, namely the norms connected to the ideal of public reason. The ideal of public reason requires that citizens are capable and disposed to justify their proposals for the arrangement of the basic structure of society to every other reasonable citizen. In other words, citizens must possess the civic virtue (CV) to use in political debates what Rawls called the method of (full) reflective equilibrium (RE). Thus, at least one reason why some features of social media are detrimental to liberal democracies is their negative impact on this important CV.

With this explanation at hand, this paper aims to demonstrate why RE as a CV is a necessary condition to foster healthy epistemic participation among citizens of liberal democratic societies. The paper concludes by critically discussing the need for change to the design or regulation of social media platforms in order to mitigate the current negative effects impacting the CV of RE in public debates.

Session 3D: Conflict, Compromise or Consensus? - Social Media and Democracy

Andreas Spahn

Social Media technologies have become a central element of the public sphere. Many scholars have argued that the rise of social media is a significant change of the public sphere that offers both chances and dangers for modern democracy (Papacharissi 2002; Fuchs 2014; Burgess and Green 2018). Recently worries have been raised about the disruptive potential of these technologies: social media platforms are leading to an increasing polarization within societies and a radicalization of individuals (Bail 2021). Some scholars are worried that we are moving towards a post-truth world (McIntyre 2018), in which it becomes more and more difficult to have a societal consensus on basic facts. At the same time classical media are losing their influence and find it increasingly difficult to survive in the market competition.

The way in which people communicate and deliberate about their ideas, however, plays a central role in the formation of public opinion within democracies. Democracy theorists have long argued that the public sphere plays a vitally important role within a democracy (Habermas 2010). In this paper we revisit discourse ethics to investigate whether it can offer orientation for the current changes within the communication in the public sphere (Habermas 1987; 1993; Apel 2016; Fuchs 2021). Discourse ethics has, however, often been criticized as being overly idealistic: often a consensus in a discourse is not possible, and maybe not even desirable (Mouffe and Holdengraber 1989). What is needed therefore is a broader framework that does not only paint one unrealistic ideal of deliberation, but distinguishes between different modes of communication and their internal moral principles.

With some exaggeration one could muse that social media might have been specifically designed to negate all four validity claims of communication, that Habermas identified: With regard to intersubjectivity they do not seemed to foster consensus, but spread conflict and polarization; with regard to objectivity they are not aimed at truth, but rather suitable for misinformation and fake news; with regard to subjectivity they are not designed for authentic and sincere communication, since individuals can hide behind anonymous identities.

The presentation will apply key insights of discourse ethics, particularly the theory of communicative validity claims, to social media to help guide discussions and deliberations on these platforms. We suggest to distinguish different communicative situations (i.e. 'conflict', 'compromise' and 'consensus') and investigate the role of the validity claims for each of these communicative situations.

References

Apel, Karl-Otto. 2016. Diskurs und Verantwortung. Frankfurt am Main: Suhrkamp.
Bail, Christopher A. 2021. Breaking the Social Media Prism. Princeton University Press.
Burgess, Jean, and Joshua Green. 2018. YouTube: Online Video and Participatory Culture. John Wiley & Sons.
Fuchs, Christian. 2014. 'Social Media and the Public Sphere'. TripleC: Communication, Capitalism & Critique (1): 57–101. https://doi.org/10.31269/triplec.v12i1.552.
———. 2021. 'The Digital Commons and the Digital Public Sphere'. Westminster Papers in Communication and Culture 16 (1). https://doi.org/10.16997/wpcc.917.
Habermas, Jürgen. 1987. Theory of Communicative Action: Lifeworld and System. Beacon Press.
———. 1993. Justification and Application : Remarks on Discourse Ethics. Cambridge, Mass.: MIT Press.
———. 2010. Strukturwandel Der Öffentlichkeit. Frankfurt am Main: Suhrkamp.
McIntyre, Lee C. 2018. Post-Truth. Cambridge, MA: MIT Press.
Mouffe, Chantal, and Paul Holdengraber. 1989. 'Radical Democracy: Modern or Postmodern?' Social Text, no. 21: 31. https://doi.org/10.2307/827807.
Papacharissi, Zizi. 2002. 'The Virtual Sphere: The Internet as a Public Sphere'. New Media & Society 4 (1): 9–27. https://doi.org/10.1177/14614440222226244.

Session 3D: A Pragmatic Method for Engineering Concepts and Technologies

Irene Olivero

The introduction and increasing use of Artificial Intelligence (AI) in several areas of our lives has brought significant changes and challenges. AI is the common denominator of several so-called "socially disruptive technologies" (SDTs) – e.g., robotic friends, virtual assistants, deepfakes, self-driving vehicles, etc. Socially disruptive technologies bring about ethical and social challenges and problems; some of these technologies give rise to so many issues that we need to consider whether it would be more beneficial to get rid of them or redesign some of their components. Here is where the recent trend of conceptual engineering (CE) comes in handy. I argue that Amie L. Thomasson's (2020) method for conceptual engineering may represent a fine methodology to guide the way we should "reshape" our socially disruptive technologies. Thomasson advocates for adopting a pragmatic approach to CE that takes "the function of our (ranges of) concepts as playing a central role (2020, 440). There are three crucial passages in this method. We should: employ reverse engineering, i.e., look at the genealogy of the concept under examination (cf. Plunkett 2016) and try to determine what it does and can do; identify the function(s) (if any) the concept in question should serve and is to serve given the goals and purposes we have; finally, constructively engineer the concept at stake, given the function(s) we need it to perform (cf. also Haslanger 2000). I here argue that this method may be of inspiration for the philosophy of technology when it comes to "engineering" the ontology and metaphysics of the emerging SDTs, for it can also take into account the ethical and social concerns these technologies bring about. For example, using this method, one may ask: what is the proper function (if any) of deepfakes? What function(s) (if any) do we want deepfakes to serve? Answering these questions may help us decide how to re-engineer our emerging SDTs (or whether to eliminate some of them, e.g., were they not to serve the function we want them to). As we do with concepts, asking ourselves what function SDTs serve and what function we want them to serve may be the key to understanding what we should keep about these entities and, in some cases, even whether we should keep them altogether. By applying this method, we shift the focus to the normative aspect, to thinking about what function(s) our SDTs should play. The importance of this method is evident in the fact that we would re-evaluate our SDTs a posteriori, i.e. after we come to know possible harmful uses not foreseen when we conceived them.

Session 3E: Understanding sustainability as a changing value in lithium battery technologies

Yunxuan Miao

Interpretations of ethical values reflect their social and historical environment. As such, while batteries are powering the massive amounts of electronics in modern society, views on what is "good" have varied temporally within the design of those batteries. When battery technologies may involve many social and ethical values, the value of sustainability can serve as a starting point, since sustainable development always requires a balance of economic growth, environmental conservation, and social equity (Gomis et al. 2011; Oden 2016). Therefore, this study explores the patterns of value change by analysing the value of sustainability in the development of lithium battery (LiB) technologies.

Through topic modeling and literature review, this study first examines the shift of focus with regards to sustainability in LiB technology research. In general, results show that the interpretation of sustainability has been gradually extended from merely focusing on LiB's service lives to their full life cycle, while the relative weights of different sustainability issues have changed over time. For instance, regarding social sustainability, most researchers merely focus on possible hazards for consumers/users in the last century, but recently increasing attention has been paid to social problems in mining, production and recycling.

In discussion of the results, an explanation is provided in terms of a historical perspective by analyzing the reverse salients in the sociotechnical system of LiBs. It is worth noting that the issue of high flammability remerges when LiBs' production has risen and become more diversified. Besides, new applications of LiBs, including electric vehicles and energy storage systems, have led to a rapidly growing demand for LiB production, triggering some reverse salients such as increasing cost and shortage of raw materials, which may have a negative impact on their economic and environmental sustainability. Moreover, this study considers the connections between reverse salients and interpretations of sustainability — a reverse salient may emerge when the balance of sustainability is upset — which could facilitate our understanding of sustainability as a changing value.

In conclusion, implications for sustainable development are considered based on previous discussions. This study, with a focus on research papers, only presents how the meaning of sustainability has changed in academic communities. To better understand value change in socio-technical systems, issues such as how value conceptualizations have changed in different social groups may need to be addressed.

Session 3E: Direct Potable Water Reuse – Change inducing for moral values?

Karen Moesker

This research is the first attempt to unpack the philosophical implications of Direct Potable Water Reuse (DPR) on the relationship between humans and nature that is realized through water. DPR separates water and nature, potentially disrupting existing physical and symbolic meanings connected to water. Consequently, DPR may bring about the need to distinguish between two sorts of water – the "natural" and the "artificial". While, "natural" water is often seen as the source of life, whereas "artificial" water resides in a circular technological system that exists independently of its natural environment. This gives rise to normative uncertainties when assessing the acceptability of DPR.

DPR disrupts our understanding of the human-nature relationship because it removes an essential part of the drinking water provision cycle: the need for nature. In its essence, DPR relies on current centralized water systems and aims to connect the wastewater with the drinking water stream by adding additional treatment steps in between, rendering it circular (Voulvoulis, 2018). The majority of the employed water treatment technologies are well established and form independent integral parts of our built environment (see, i.e.: Gerrity, Pecson, Trussell, & Trussell, 2013). Consequently, DPR does not implement new (singular) technology. Rather it is a socio-technical system that assembles existing technologies in a novel way. The disruptiveness, thus, lies in the assembly, not the technology itself.

Before the introduction of DPR, the considered water treatment technologies raised large controversies because they do not contest the understanding of the human-water relationship. Waste and drinking water treatment technologies are placed between nature and humans. As such, these technologies may increase the distance of humans to nature but drinking water is still retrieved from nature. These technologies can thus be seen as enhancing the natural capacity of providing humans with drinking water.

DPR on the other hand changes the source of the drinking water. When utilizing DPR, water is no longer retrieved or dumped in nature where it is then (miraculously) purified. Instead, water is kept and circulated within the technical water system. The new assembly of uncontroversial technologies removed an essential part of traditional water systems: nature. The disruptiveness of DPR is thus not the introduction of new technology but its potential to replace nature as a water source. This replacing aspect of DPR disrupts the human-nature relationship that we have through drinking water. It creates uncertainty about how we value water and how we relate to nature through it. This research tries to deepen our understanding of whether water can be artificial and its philosophical meaning. Deeper insights into the current normative uncertainties may then contribute to understanding and addressing current public acceptability issues regarding this technology.

Session 3E: Women in the Ever Evolving Marine Robotics

Simona Aracri, Rosangela Barcaro & Massimo Caccia

International organisations, like the UN and EU, are striving to increase awareness and take measures for eliminating stereotypes and bridging the gender gap [1]. The role of technology is constantly evolving in our society and, more importantly, in the Blue Economy, which plays a key role in the Sustainable Development goals, e.g. resource management, development of new and inclusive technological education (quality education) and gender equality [2]. The call for an autonomous, smart and ever connected management of the resources of the planet is pushing the boundaries of Earth Observations, industry management and health systems. In particular, research and development in marine robotics is evolving quickly. This research branch is a kaleidoscope of diversity as far as applications, tech and innovation are concerned. In recent years, the topic of gender diversity and inclusivity within marine robotics received increasing attention. The EU included the sector “marine robotics” in the EMFAF-2022-PIA-WBE - Women in the Blue Economy call. The Institute of Marine Engineering of the National Research Council of Italy plays an important role gathering data and creating a network to raise awareness among researchers and technologists. Since 2020, the institute has worked to give visibility to the gender issue in marine robotics. In the framework of the Blue RoSES [3] project, it created a community that celebrates success stories of women in the field. Such a community worked to stimulate discussion and enhance the sensitivity of the colleagues through media [4] platforms and in person events in marine robotics conferences, e.g. Breaking the Surface field workshop [5]. The group explores the gender panorama among high-profile researchers and technologists, to better understand the gender stereotypes and gaps still existing in marine robotics and beyond. Recently, within the MONUSEN [6] project, the group also developed a quiz [7] to involve researchers in a fruitful discussion about the role of gender minorities in the thriving marine robotics community. Our presentation will expand upon the issues that emerged during the discussion.

References

[1] R J Shellock, C Cvitanovic, M C McKinnon, M Mackay, I E van Putten, J Blythe, R Kelly, P Tuohy, K M Maltby, S Mynott, N Simmonds, M Bailey, A Begossi, B Crona, K A Fakoya, B P Ferreira, A J G Ferrer, K Frangoudes, J Gobin, H C Goh, P Haapasaari, B D Hardesty, V Häussermann, K Hoareau, A-K Hornidge, M Isaacs, M Kraan, Y Li, M Liu, P F M Lopes, M Mlakar, T H Morrison, H A Oxenford, G Pecl, J Penca, C Robinson, S A Selim, M Skern-Mauritzen, K Soejima, D Soto, A K Spalding, A Vadrot, N Vaidianu, M Webber, M S Wisz, Building leaders for the UN Ocean Science Decade: a guide to supporting early career women researchers within academic marine research institutions, ICES Journal of Marine Science, Volume 80, Issue 1, January 2023, Pages 56–75, https://doi.org/10.1093/icesjms/fsac214
[2] https://en.unesco.org/sustainabledevelopmentgoals
[3] http://www.blueroses.eu
[4] womeninblue.eu
[5] R. Lupu, M. Caccia, E. Zereik and R. Barcaro, "Women in Blue: Toward a Better Understanding of the Gender Gap in Marine Robotics [Women in Engineering]," in IEEE Robotics & Automation Magazine, vol. 29, no. 4, pp. 138-140, Dec. 2022, doi: 10.1109/MRA.2022.3213467.
[6] http://www.monusen.ucg.ac.me/
[7] https://drive.google.com/drive/folders/1DWXGO7GOWBN15P9eLKBGeZRGEQlwRO3t?usp=share_link

Session 4A: Designing philosophies of education as fruitful bridges between philosophical ideas of engineering and realities in engineering education

Daniel Marom

A significant practical contribution that philosophy can make to engineering is through the articulation of philosophies of engineering education and their translation to the language of curricular and pedagogical theory and practice in the training of engineers. While such philosophies of education draw upon profound ideas in the philosophy of engineering, they also provide the basis upon which corresponding curriculum and subject matter, student selection and induction processes, modes of non-clinical and clinical teaching and learning, faculty training, and evaluation can be designed and implemented at every stage of engineering education.

The potential for such impact depends on the effort being guided by rigorous methodological guidelines connecting philosophical ideas with educational practice. Such a methodology was developed by the philosopher of science and education, Israel Scheffler and the curriculum expert Seymour Fox and was employed by Daniel Marom in efforts to contribute to the improvement of professional training in education, medicine, psychotherapy, and spiritual counselling. This methodology focuses on facilitating articulations in and between specific domains on a five-level dialectical continuum between philosophy and educational practice (Fox, Scheffler & Marom, 2003; Marom 2009/2022).

The proposed workshop will begin with an introduction to this methodology - its rationale, its five-level theory-practice continuum, and practical guidelines for movements in and between the various levels. Examples of the application of these guidelines will be presented from various professions.

This introduction will be followed by exercises in educational planning, implementation and evaluation in which workshop participants will rehearse movements in and between specific levels within the context of developing training activities for engineering. For example, participants might develop a portrait of the ideal candidate for an engineering program based on a systems philosophy of engineering (Usman and Bell, 2021) and a corresponding mode of selection interview. Or they might design an in-service mentoring pedagogy in transport engineering guided by social ideas relating to distributive justice (Martens, 2012).

Finally, the workshop will conclude with an open discussion exploring and appraising needs and avenues for further development of philosophies of engineering education and their translations to practice in diverse settings. One current initiative focused on the articulation of an international compendium of alternative visions of engineering education and its teaching to engineering education leaders will be presented as a reference point for discussion.

References

Seymour Fox, Israel Scheffler and Daniel Marom (eds.), Visions of Jewish Education, edited (New York: Cambridge University Press: 2003).
Daniel Marom: “Four Lessons I Learned from Seymour Fox on Education” (2022 translation from Hebrew version published in 2009): https://www.researchgate.net/publication/358743414_Four_Lessons_I_Learned_from_Seymour_Fox_on_Education. A short description of the five-levels theory appears on pages 44-49.
Karl Martens, “Justice in transport as justice in accessibility: applying Walzer’s ‘Spheres of Justice’ to the transport sector,” Transportation 2012: DOI 10.1007/s11116-012-9388-7.
Akeel Usman and Sarah Bell, “Systems Engineering as Engineering Philosophy” in The Routledge Handbook of the Philosophy of Engineering, edited by Diane P. Michelfelder and Nilke Doorn (New York: Routlege, 2021) 176-190.

Session 4B: How do engineers model engineering work? Contrasting cost and schedule models to understand modeling practices in complex system developments

Zachary Pirtle

In complex systems developments, engineers often need programmatic analysis models to help manage and make design and integration decisions. Programmatic analysis models assess cost, schedule, and risk for a development program, helping to predict a future project’s cost based on factors like past history, current judgments of technical complexity, and understanding the planned scope of work by engineers and technicians. Programmatic analysis models range from parametric cost and schedule models, quantitative risk models, cash flow models, and joint cost and schedule confidence level models. These models can strongly impact the engineering work – perhaps leading a manager to descope technical content to stay within budget – and can shape engineering policy, by guiding decision makers on what types of systems should be built for expected costs.

The practical impact of these assessment models – perhaps especially including when they fail to predict future cost and schedule overruns – calls for a deeper conceptual understanding of how they work. How does the way engineers represent a development project in a model shape the expected cost and schedule of that project? Are there implications of this for engineers and managers? Programmatic analysis models can bear a surprisingly unclear relation to actual engineering projects, representing the work of hundreds of engineers and technicians through a series of idealizations that allow the models to predict cost, schedule and risk.

To explore these nuances, I compare two functionally similar programmatic analysis models (Cash Flow (CF) and Joint cost and schedule Confidence Level (JCL)) models that model the same phenomena but do so using different approaches. Using a synthesized framework based on an extensive survey of systems engineering and philosophy of modeling literatures, I contrast the models across several dimensions of model composition, including causal idealization, parameters, data inputs, and modeler skillsets. While the create similar outputs, these models differ significantly in terms of the way they idealize engineering work, respectively describing the work of an engineer or technician as either cost or time. Using a series of structured thought experiments, I'll describe how differences like these matter significantly, shaping the ways in which a model can identify programmatic challenges as well as how they can respond to them. I will highlight how indirect ways in which a model is built can shape analysis results, such as by limiting the scope of how much managers can actually manage a problem.

I'll discuss several implications of this work. This process for describing how specific models are independent from one another can be used in other areas of engineering modeling, with the process helping to understand how model composition shapes insights about real-world systems. This conceptual work also can help engineering practitioners in designing, selecting and using models based on their contrasting strengths and weaknesses. For programmatic assessments, these results can help focus debates on what uncertainties matter for complex systems developments, to help explore how to best enable affordable and reliable developments.

Session 4B: Epistemic achievements of engineers in relation to sociotechnical systems: From technological knowledge to engineering understanding

Vivek Kant & Michael Poznic

The epistemic achievements of engineers are usually discussed in terms of knowledge. Different types of engineering or technological knowledge are analyzed in the literature (Vincenti 1990, De Vries 2003, Houkes 2006, Meijers & Kroes 2013, Norström 2015). This paper argues for the recognition of an alternative epistemic state of engineers beside knowledge, i.e., understanding (Elgin 2017). Especially in cases of complex objects of investigation, engineers or other practitioners need to bring different pieces of knowledge together, and an insight into the individual pieces, but also the relations among the pieces, and the overall body of information is crucial. Such a demanding epistemic state of understanding is particularly important when it is the task to design or shape a sociotechnical system. Such systems are hybrids of people and technologies in various modalities, constrained by regulations and various forms of governance. In this milieu, the engineer is one entity, albeit an important one, amongst several entities that deal with the creation, operation, and functioning of the system. Since these systems involve a number of engineering disciplines, there is a need for recognizing different aspects of engineering knowledge. So, the discussed types of knowledge in the philosophical literature are indeed important. This paper is not meant to criticize the contributions but to broaden the horizon of established discussions. In addition to the task of analyzing different types of knowledge, a related challenge is to spell out what the engineer has to do to integrate particular pieces of knowledge into a coherent whole. While individual engineers require specialized knowledge, conduct of work in these systems also requires a generalized engineering understanding of the functioning of the overall system, or large parts of it. This generalized understanding may exist in degrees and could be representing the functional, the structural as well as other aspects of the system such as social or moral aspects among others. The key idea to emphasize here is that epistemic achievements in sociotechnical systems require a discussion of both engineering knowledge and a generalized engineering understanding. We will outline the notion of engineering understanding and its scope with a few examples of its applicability. We will also highlight that this concept has not received the same explication as engineering knowledge and is required for comprehending epistemic achievements of engineers in complex sociotechnical systems.

References

de Vries, M. J. (2003). The Nature of Technological Knowledge. Techné: Research in Philosophy and Technology, 6(3), 117–130.
Elgin, C. Z. (2017). True Enough. MIT Press.
Houkes, W. (2006). Knowledge of artefact functions. Studies in History and Philosophy of Science, 37(1), 118–131.
Meijers, A., & Kroes, P. (2013). Extending the Scope of the Theory of Knowledge. In M. J. de Vries, S. O. Hansson, & A. Meijers (Eds.), Norms in Technology. Springer.
Norström, P. (2015). Knowing how, knowing that, knowing technology. Philosophy and Technology, 28(4), 553–565.
Vincenti, W. G. (1990). What Engineers Know and How They Know It. The Johns Hopkins University Press.

Session 4B: Epistemic uncertainties in flood frequency analyses and the ethics of belief

Franscisco Olivera & Martin Peterson

In a well-known paper published in 1877, Clifford James coined the term "ethics of belief", which refers to the norms that should guide our belief-forming habits. In the debate on the ethics of belief, several practically useful epistemic tools and distinctions have been developed. In this paper, we apply ideas from the literature on the ethics of belief to two approaches to flood frequency analysis. The goal of flood frequency analysis is to estimate the probability that a given flow rate in a particular river (at a particular point) will be exceed in any given year. What is, for instance, the probability that the flow rate of the Rhine River in Lobith will exceed 6000 cubic meters per second at any point in any given year? Calculating this probability is key for designing hydraulic structures, developing flood mitigation plans, urban planning, and so on. However, the probability that the flow rate will exceed a given value can be calculated in several different ways. In this talk, we are going to discuss two. According to the first approach, which we will refer to as the traditional approach, the parameters of the probability distribution are estimated from the annual maximum instantaneous flow rates in the river. Typically, the water resources engineer will have access to historical data for a particular flow gauging station in the river and use the annual maxima time series (which is recommended to span more than thirty years). This approach relies crucially on the assumption that the time series is stationary (i.e., its statistics do not change over time) and that the observations are statistically independent. However, both these assumptions are questionable. Because of climate change, the statistics of the time series are almost certainly undergoing continuous change. Moreover, observed wet and dry multi-year periods in flow records, suggest that the observations are not independent. In order to overcome these and other problems with the traditional approach, an alternative, "high resolution" approach has been proposed. The high-resolution approach is based on the idea that using more data is more desirable than using less: instead of basing the estimate on annual maxima, the high-resolution approach uses all available values in a year. Obviously, this makes the independence assumption even more unrealistic. Epistemic uncertainties induced by climate change remain unresolved. However, a key advantage of the high-resolution approach is that it is going to do more epistemic justice to years in in which there is more than one extreme flood event.
We are open to the idea that there is no objective truth (in a sense that we will not attempt to define) about what the probability that a given flow rate in a particular river at a particular point is going to be exceeded is. The answer clearly depends on epistemic choices made by the water resources engineer, and such choices need to be based on an ethics of belief. In the talk, we sketch two approaches to such an ethics of belief for water resource engineering: a consequentialist ethics of belief, and a virtue-based ethics of belief. We do not take a stand on which approach is best. Our aim is just to draw attention to the fact that conclusions about something that seems to be a purely scientific question ("What is the probability that a given flow rate in a particular river, at a particular point, is going to be exceed in any given year?") depends crucially on the engineer's ethics of belief.

References

Clifford, W.K., 1877 [1999], "The ethics of belief", in T. Madigan, (ed.), The ethics of belief and other essays, Amherst, MA: Prometheus, 70–96.
Naess, A., O. Gaidai and O. Karpa, 2013, "Estimation of Extreme Values by the Average Conditional Exceedance Rate Method", Journal of Probability and Statistics, Volume 2013, Article ID 797014, http://dx.doi.org/10.1155/2013/797014.

Session 4C: A Frankian Problem with Genetic Engineering

Brandon Long

Genetic enhancement has been subjected to many criticisms, one of which is that enhancement aimed at positional goods is obviously fruitless. This has been accepted with a bit too little skepticism, and an examination of the nature of behavioral traits demonstrates how this is so. Traits are heritable, and even non-biological traits such as socioeconomic status are heritable. Other biological traits such as disagreeableness likely contribute to the heritability of socioeconomic status. Disagreeableness is mostly a positional and extrinsic good because, it is theorized, being disagreeable means a higher likelihood of swaying workplace managers to give you a raise over others. The problem being that if disagreeableness is targeted for genetic enhancement, everyone will be disagreeable, and yet no net gain can be achieved if everyone has access to enhancement technologies. Worse yet, disagreeableness comes at cooperation costs in the workplace, making this outcome worse than if no one had enhanced their offspring. This is a convincing argument against enhancement of positional goods. However, for people who are very agreeable or disagreeable, acting otherwise is quite difficult. This means that achieving virtue is difficult when one must find the mean act in any given context, when one is unlikely to act a certain way. Bending your behavior to the mean through deliberation is key for Aristotelian virtue, as the struggle to act in all the proper ways is part of the virtuous act. The enhancements proposed here avoid much concern from anti-enhancement arguments that state that enhancement makes virtuous behavior so easy as to devalue it, and undermines autonomy. Enhancements that move offspring behavior closer to the mean from the extreme end of a bell curve for a trait give offspring the best ability to deliberate, without making their behavior predictable and therefore lack autonomy. In fact, this edit makes behavior less predictable. Offspring autonomy can not be undermined when enhancements give better deliberative and less predictable behavior to offspring. Thus, these edits moving an offspring from the extreme end of a bell curve for a trait closer to the mean will not be subject to these anti-enhancement arguments.

Session 4C: Qualification and quantification of fairness for diversity-aware and inclusive mobility

Camilla Quaresmini, Eugenia Villa, Valentina Breschi, Viola Schiaffonati & Mara Tanelli

Policies promoting alternative mobility solutions finalized at reducing carbon emissions seldom account for social factors, which have been proven crucial for the individual adoption of new technology-enabled behaviours. This often undermines their effectiveness, producing the unwanted effect of amplifying differences at the edges of society. While the quantitative approaches usually exploited for policy design ignore the epistemic charge of technical assumptions, without pondering their epistemological presuppositions, purely qualitative methods risk being general, bringing the issue to a high level of abstraction.

Our solution is thus to combine conceptual tools from philosophy with engineering, making technology an active mean to implement social justice and reduce inequalities. By mathematically modeling and quantifying fairness, our goal is to design diversity-aware and inclusive technologies in the context of sustainable mobility. Hence, we want to propose an approach based on data collected through surveys to embed socio-economic factors in the design of new mobility strategies, that quantitatively account for social fairness in a control-oriented and dynamic fashion. By representing individuals as agents in a network, and assuming that the diffusion of new mobility habits is driven by personal attitudes and mutual influences, we develop optimal control techniques to design fostering policies that minimize costs, while rewarding social justice. We formalize the concept of fairness and embed it into the problem via two elements: equality, i.e., attaining an even distribution of the available resources, and equity, namely promoting agents to be comparably close to the final target of adoption, in spite of their initial different attitudes. The proposed framework thus allows us to iteratively integrate fairness in the policy design process, by enabling its dynamic reassessment.

Combining conceptual investigation and empirical study [1], this approach of interdisciplinary coexistence allows us to consider quantitative aspects, while characterizing them philosophically. This is the case of the homophily principle, i.e., preferential in-group interactions, which should be taken into account when defining agents' connections, being one of the most robust empirical regularities of social life. However, leading to disproportionate in-group interaction, exploiting homophily only in policy design can be considered as reflecting some degree of discrimination [2], involving epistemic exclusions. Therefore, a meta-reflection on the impact of adopted perspectives becomes necessary, in the general context of a responsible policy optimization which must involve philosophical evaluations, going beyond standard measurements.

References

[1] Batya Friedman and David Hendry. Value Sensitive Design: Shaping Technology with Moral Imagination. 2019. [2] Cailin O'Connor. The Origins of Unfairness: Social Categories and Cultural Evolution. Oxford University Press, 2019.

Session 4C: Towards a postphenomenology of mutual recognition in AI design

Fernando Secomandi

In this presentation, I will advance an ethical framework for the design of liberatory user interfaces for AI technologies. I will revise Verbeek's (2011) postphenomenological framework of mediated morality by incorporating the Hegelian concept of mutual recognition (Hegel, 2018; Williams, 1998). I will show how Verbeek's framework falls short in explaining the intersubjective constitution of human-world relations and argue that users' freedom in relation to AI can only be attained by mediation of an other human intentionality. In the process, I will comment on other contributions within the philosophy of technology that have called for critical expansions of postphenomenology towards the social-political domain (Bantwal Rao et al., 2015; Botin et al., 2020; Feenberg, 2009; Kaplan, 2009) and that have advanced Hegelian approaches to human-robot relations (Coeckelbergh, 2015; Gertz, 2018; Waelen, 2022).

References

Bantwal Rao, M., Jongerden, J., Lemmens, P., & Ruivenkamp, G. (2015). Technological Mediation and Power: Postphenomenology, Critical Theory, and Autonomist Marxism. Philosophy & Technology, 28(3), 449–474. https://doi.org/10.1007/s13347-015-0190-2

Botin, L., de Boer, B., & Børsen, T. (2020). Technology In Between the Individual and the Political: Postphenomenology and Critical Constructivism. Techné: Research in Philosophy and Technology, 24(1/2), 1–14. https://doi.org/10.5840/techne2020241

Coeckelbergh, M. (2015). The tragedy of the master: Automation, vulnerability, and distance. Ethics and Information Technology, 17(3), 219–229. https://doi.org/10.1007/s10676-015-9377-6

Feenberg, A. (2009). Peter-Paul Verbeek: Review of What Things Do. Human Studies, 32(2), 225–228. https://doi.org/10.1007/s10746-009-9115-3

Gertz, N. (2018). Hegel, the Struggle for Recognition, and Robots. Techné: Research in Philosophy and Technology, 22(2), 138–157. https://doi.org/10.5840/techne201832080

Hegel, G. W. F. (2018). Georg Wilhelm Friedrich Hegel: The Phenomenology of Spirit (T. Pinkard & M. Baur, Eds. & Trans.). Cambridge University Press. https://doi.org/10.1017/9781139050494

Kaplan, D. (2009). What Things Still Don't Do. Human Studies, 32(2), 229–240. https://doi.org/10.1007/s10746-009-9116-2

Verbeek, P.-P. (2011). Moralizing Technology: Understanding and Designing the Morality of Things. University of Chicago Press.

Waelen, R. A. (2022). The struggle for recognition in the age of facial recognition technology. AI and Ethics. https://doi.org/10.1007/s43681-022-00146-8

Williams, R. R. (1998). Hegel's Ethics of Recognition. University of California Press.

Session 4D: The Normative Need for Trust in AI

Sara Blanco

Trust is often considered the glue of a healthy society. We need others and others need us. Trusting puts the one who trusts in a vulnerable place. Because of this, trust is generally placed only in those to whom we believe we can delegate something. Both the general public and a part of academia often talk about trust and distrust in artificial intelligence (AI). However, it is unclear whether the concept of trust can suitably be extended to non-human entities. In this paper, I argue that trust relationships imply moral responsibility and it is precisely this implication that makes trust a normative goal for human-AI relationships.

AI systems are not technical artifacts like any other. AI is a novel technology able to achieve results without being explicitly programmed on how to reach such results. Instead, AI 'learns' from big volumes of data and comes out with its own paths towards outcomes. Thus, how we use and relate to AI differs from how we relate to previous technology. At the moment, AI models are used in a variety of domains, from finance to healthcare. Due to its potential to affect people's lives, AI is widely considered a socio-technical tool. This means that its successful implementation concerns the interaction of social and technical factors. Then, conceptualising the kind of relationship that we, as a society, aim to have with AI is crucial for the successful implementation of the latter.

It is often argued that technology cannot be trusted but, at its best, only relied on (Ryan 2020). This view puts the focus on technical success to determine whether a system oughts to be implemented; that is, whether the predictions made by the system are mostly accurate. However, in domains such as medical diagnosis, a successful outcome is not just a technical solution to a technical problem. Moral responsibility should be taken into account, being an important fact to consider in the use and implementation of AI. Because of this, I claim that aiming for reliance overlooks the role of AI in society. Therefore, reliance conceptually fails to capture the relationship that humans ought to have with artificial intelligent systems.

I understand trust as a relational concept that implies moral responsibility. Trust refers to relationships in which a trustor willingly makes themselves vulnerable towards a trustee and accepts the risk of being betrayed. I argue that willingly making oneself vulnerable towards someone implies placing moral responsibility on that someone. From the trustee's side, committing to not abuse others' vulnerability implies assuming moral responsibility. Because of the contexts in which AI can be used, this kind of responsibility attribution becomes normatively necessary. Therefore, trust is a suitable goal to aim for when shaping human-AI relationships.

The novelty of AI creates a conceptual gap between this kind of system and other technological artifacts. Given the kind of situations in which AI can be used, I defend that technical success is not enough to satisfy the normative goals of AI. Rather than a drawback, I propose the moral implications of trust as the reason to normatively aim for trust in human-AI relationships.

Session 4D: AI as Inventor? A patent system's call for philosophy

Jens Pilger & Michael W. Schmidt

Increasingly AI systems take over tasks that previously mainly relied on creative human engagement. Take for example current AI systems, trained on big data, which generate pictures based on text input. If humans would have generated these pictures, they might be considered artists and hold corresponding intellectual property rights. Thus, the question arises what is the adequate status of the AI system in such a case and what statuses to assign to the persons involved in the process, such as the humans that used the AI system, the ones that owned the AI system or created it and those whose works were part of the training data.

The same question arises in the context of technological innovation. Should AI systems, which generate something that would be seen as an invention, were it created by humans, considered as the inventors? And if so, who would be entitled to hold corresponding patent rights? Alternatively, should there be no right to a patent granted in such cases?

This is not only an academic question as the DABUS case shows, which recently challenged the patent system worldwide. DABUS, an AI system, was designated as inventor in patent application proceedings in several jurisdictions, including the Patent Offices of the UK, the USA, Germany, South Korea, South Africa, and the European Patent Office. With the exception of South Africa all applications were rejected based on the reason that the designated inventor must be human (George & Walsh, 2022; Drexl et al., 2021; Buzu, 2022). At this point, the legal situation seems to be clarified in these economically important countries: AI-generated inventions are in principle patentable, but only a human can be designated as the inventor in the patent documents.

However, this legal status quo can and should be scrutinized from a more philosophical position: Is it adequate to grant a human the status as an inventor where the essential part of the innovation was done by an AI system? In addition and even more challenging: which human being should be designated, and on which basis, as the inventor? Moreover, is it adequate to reject patents where no human inventor can reasonably be designated? These questions will become more pressing with the further development and implementation of AI systems, and the question arises if, and how, the patent system could be adapted to the new challenges coming from AI.

We draw on a theory of patents, which highlights the social function of patents, in order to find preliminary answers to these questions.

References

Buzu, I. (2022). The DABUS Affair. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4175009
Drexl, J., Hilty, R., Kim, D., & Slowinski, P. R. (2021). Artificial Intelligence Systems as Inventors? A Position Statement of 7 September 2021 in View of the Evolving Case-Law Worldwide. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3919588
George, A., & Walsh, T. (2022). Artificial intelligence is breaking patent law. Nature, 605(7911), 616–618. https://doi.org/10.1038/d41586-022-01391-x

Session 4D: Manipulative technology

Michael Klenk

Interpersonal manipulation is, for better or worse, a staple of human social life. While scholarship on the concept and ethics of manipulation has laid dormant for many years, it has recently been rejuvenated by growing concerns about (online) manipulation through technology (cf. Jongepier & Klenk 2022).

The predominant focus in the recent debate about manipulation and technology has been on whether and how technology affords greater, or more significant, opportunities for humans to manipulate other humans (e.g. Hancock et al 2021). The question of whether and how technology itself can be manipulative has received far less attention. This is a shortcoming, especially in light of some of the key concerns in the ethics of technology with value-laden technology (cf Winner 1980, Klenk 2021).

In this talk, however, I attempt to remedy that shortcoming. I will argue that, based on an appropriate understanding of the concept of manipulation, it is reasonable to attribute the concept to technology even if technology lacks agential features.

First, I distinguish the predominant rights-based conception of manipulation from an interesting value-based conception of manipulation. According to the latter conception, manipulation only requires purposeful but not intentional influence. Second, I argue that technological artefacts can be said to exhibit purposeful influence on users. Ergo, technology itself can be a source of manipulate. Third, I argue that the manipulativeness of any given technology depends on whether it hinders the human aim of deliberation in specific ways.

In the final part of the talk, I consider the (dis-)value of manipulation by drawing on (socio-)epistemological insights into the social source of epistemic goods that manipulative technologies can obstruct.

Session 4E: What should we strive for in engineering ethics education? A complex and granular mapping of Engineering Ethics Education competencies and skills

Lavinia Marin

This presentation proposes a reassessment of the goals of engineering ethics education based on an alternative way of conceptualizing these goals that is more granular and hence more achievable in educational settings.

In engineering ethics education (EEE), several moral competencies and skills are usually posited as the goals for students to acquire at the end of the ethics courses. Such competencies are Moral sensibility, Moral analysis skills, Moral creativity, Moral judgment skills, Moral decision-making skills, Moral argumentation skills, etc. (van de Poel & Royakkers, 2011, p. 2). However, a problem of scope emerges. In educational theory, learning goals are divided by levels of abstract cognitive skills, making them more granular. Thus, according to Blooms's taxonomy, learning goals should be divided into skills that pertain to remembering, understanding, applying, analysing, evaluating and creating (Bloom, 1956). The EEE learning goals seem cognitively complex, but it is up to the instructors to decide at which Blook level they want to situate their teaching and goals. The goals of EEE are ambitious because they can cover several Bloom levels at once, making them difficult to achieve in practice. In the first part of the presentation, I will argue that the EEE competencies and skills can be further divided into a more granular mapping to cover the Bloom taxonomy, thus making them more precise.

In the second part, I showcase an alternative way of dividing the EEE competencies into several levels of abstraction based on Luciano Floridi's framework of information ethics (2013). Based on Floridi's framework, a system can be analysed at the agent level, patient level, relational level and ecosystem level. Each of these levels gives rise to specific questions about the ethical issues arising at that level (e.g. who is the moral agent, who is the moral patient, what kind of ethical relations are taking place, what kind of ecosystem is created with all the actors and their environment, what kind of affordances for moral action emerge or are designed?). The levels of moral analysis are complementary: one can find ethical problems at one level while at another level everything may be acceptable. I will offer an integrated mapping of the levels of abstraction for ethical analysis (Floridi's model) that breaks down the generic moral competencies of EEE into more granular skills and competencies. I will then map out these granular ethical competencies back to the Bloom levels of cognitive skills. I end the presentation with a discussion evaluating which granular competencies and skills in the new mapping are more attainable and which are hard to achieve in a classroom but still necessary, thus requiring new pedagogical approaches.

References:

Bloom, B. S. (1956). Taxonomy of educational objectives. Vol. 1:
Cognitive domain. New York: McKay, 20-24.
Floridi, L. (2013). The ethics of information (First edition). Oxford University Press. http://www.loc.gov/catdir/enhancements/fy1409/2012276183-b.html
van de Poel, I., & Royakkers, L. M. M. (2011). Ethics, Technology, and Engineering: An Introduction. Wiley-Blackwell.

Session 4E: Metaphors of Learning and Their Uses in Engineering Education: Reflections from an Engineering Educator

Katherine A. Goodman

Metaphor is deeply embedded in abstract thought [1], and metaphor often becomes a tool for learning. As Dewey argues, students must have a means of connecting new ideas to already familiar concepts [2]. For instance, many learn about electricity moving through wires using a metaphor of water flowing through pipes. Metaphors for learning itself often default to acquiring-a-thing: instructors discuss whether students "get a concept" and students refer to a course as "putting tools in my toolbox" for future use. Constructivism states that students are not merely vessels to pour knowledge into, but also describes the stages of development when children hang onto or let go of certain ideas [3]. Clearly this metaphor is useful, but it is also limiting. The metaphor of learning as acquiring-a-thing is reinforced by math and science education, in which students must often provide a single correct answer for a problem. Yet, as Sfard reminds us, a different metaphor for learning might be summarized as participating-in-an-activity, conceptualizing knowledge as an act that is typically collaborative [4]. This has pleasing alignment with engineering practice, which cares little what one knows if one cannot do anything with it. Here I propose yet another metaphor, learning as finding-and-making-paths. This view of learning moves students away from a closed "one right answer" mindset, toward one that explores a problem space and its solution set. A path-finding metaphor describes a learning process of locating and using known technical knowledge for a design, described by Vermaas et al. as "normal designing," whereas path-making would be "radical designing" [5]. The paths metaphor also suggests navigating and utilizing multiple levels of representation, an "inscriptional chain" that allow engineers to move between the more abstract (equation, model) to the more concrete (detailed design, built structure) [6]. The paths metaphor may better inform engineering educational practices and reveal insights as to the nature of engineering epistemologies.

References

[1] G. Lakoff and M. Johnson, Metaphors We Live By, 2nd ed. University of Chicago Press, 2003.
[2] J. Dewey, Experience and Education. New York: Touchstone by Simon & Schuster, 1938.
[3] J. Piaget, "Part I: Cognitive development in children: Piaget. Development and learning," Journal of Research in Science Teaching, vol. 2, no. 3, pp. 176–186, 1964, doi: 10.1002/tea.3660020306.
[4] A. Sfard, "On Two Metaphors for Learning and the Dangers of Choosing Just One," Educational Researcher, vol. 27, pp. 4–13, 1998, doi: 10.3102/0013189X027002004.
[5] P. Vermaas, P. Kroes, I. van de Poel, M. Franssen, and W. Houkes, "A Philosophy of Technology: From Technical Artefacts to Sociotechnical Systems," Synthesis Lectures on Engineers, Technology, and Society, vol. 6, no. 1, pp. 1–134, Jan. 2011, doi: 10.2200/S00321ED1V01Y201012ETS014.
[6] A. Johri, W.-M. Roth, and B. M. Olds, "The role of representations in engineering practices: Taking a turn towards inscriptions," Journal of Engineering Education, vol. 102, no. 1, pp. 2–19, 2013, doi: 10.1002/jee.20005.

Session 4E: Rethinking Engineering Professionalism through the Concept of Bildung

Anders Buch & Steen Hyldgaard Christensen

The German term Bildung refers to a socio-philosophical and pedagogical tradition born in a particular historical, geographical, cultural, linguistic and religious context. Conceived in the Christian world of the sixteenth century, it depends on a Germanic understanding of education distinct from that of its neighbours such as France or Great Britain at the time (Horlacher 2021, 129). Put by Wilhelm von Humboldt (1767-1935) at the heart of his project for Berlin University (Horlacher 2016), this concept of education has been widely disseminated in Central and Northern Europe where it has undergone many evolutions and reinterpretations (Sjöström et al 2017). It is also discussed in the US where the great universities were created following the Humboldian model. Recently, the concept of Bildung has come back to the front stage of academic research on education, science education, higher education and also engineering education (Christensen and Kolmos, 2006; Christensen et al 2007), but not everywhere in the world (Yang, 2022). While Bildung knows literal translations in several languages of Northern Europe (Scandinavia and Finland), and is used without being translated in the English-speaking world, it is neither translated nor used in its original version in the French academic literature, apart from works dealing with German-speaking authors or Franco-German comparative education (Wallenhorst 2016).

This communication questions the relevance of the concept of Bildung to think or re-think engineering education in the French-speaking area. First, we argue that it is necessary for researchers studying engineering education in an international context to develop a "global" competence, just like engineers are invited to do (Downey et al 2006). This concern for a better mutuel understanding already discussed about ethics issues for the engineering practice and training (Didier 2022) is also inspired by Wilhem von Humboldt's for whom the diversity of languages was not only a diversity of signs but of worldviews. This discussion might be useful for an audience not accustomed to the concept of Bildung, as well as for those who consider it a universal concept. Secondly, thanks to the typology of variations of the Bildung concept in international research on science education proposed by Jesper Sjöström (et al. 2017), we will try to identify possible bridges with the concepts which are used in the French speaking linguistic and cultural contexts when dealing with the issues for whom the Bildung concept is increasingly being discussed today.

References

Christensen, Steen H, Meganck, Martin, Delahousse Bernard. (2007) Occupational Bildung in Engineering Education, In. Philosophy in Engineering, Herning, Academica
Didier C. (2022) Engineering and Business Ethics: Revisiting the Higher Aims of Professionalism, In Christensen et al, (112-123)
Horlacher R. (2016) The Educated Subject and the German Concept of Bildung : A Comparative Cultural Histyory, New York, Routledge.
Horlacher, Rebekka. (2021) Chapter 4. German Educational Thought: Religion, Rationalism, Philanthropism and Bildung, In Megan Laverty and David Hansen (dir.) A History of Western Philosophy of Education, Vol 3 (edited by Tal Gilead), (125-149), Bloomsbury

Christensen, J., Henriksen L. B., Kolmos, A (Eds.) (2006), Engineering science, skills, and Bildung, Aalborg Universitetsforlag.
Sjöström Jesper, Frerichs Nadja, Zuin Vânia G. & Ingo Eilks (2017) Use of the concept of Bildung in the international science education literature, its potential, and implications for teaching and learning, Studies in Science Education, 53:2, 165-192
Wallenhorst, N. (2016). De la Bildung à la citoyenneté existentielle: Une approche franco-allemande de l'éducation des travailleurs sociaux. Le sociographe, HS n°9, 55-70.
Yang L. (2022). Student formation in higher education: a comparison and combination of Confucian xiushen (self-cultivation) and Bildung, Higher Education, 83, 1163-1180.

Session 4F: Revaluing the Public Debate: A Postphenomenological Account of Changing Values in the Digital Public Sphere

Anthony Longo

Since the 1990s, academics have observed how the internet and social media have come to provide an important space for community building for marginalised groups. Such a space has been more broadly described as a 'digital public sphere', commonly defined the communicative sphere of interaction provided by online media where citizens can gather to discuss common concerns. The theorization and assessment of this development is predominantly based on the public sphere theories of Jürgen Habermas and Hannah Arendt. Because of this, is has been widely argued that the values that social media privilege contradict traditional models of the public sphere. In the past decade, it has been well documented that social media environments privilege communication that adheres to the logic of popularity (Bucher 2018); that algorithmic curation supports the 'outrage culture' of the Internet (Gertz 2018); that social media creates a truth-less public sphere by design (Marres 2018); and that internet-specific communication features such as emojis, GIFs and memes contribute to a trolling culture (Phillips 2015). While these accounts address the fact that social media has changed the values at work in the public debate, they do not yet fully explain how this is possible.

This paper discusses two philosophical approaches that aim to explain the relationship between technological change and value change: critical constructivism and postphenomenology. Critical constructivism draws from a (neo-)Marxist approach and claims that technologies are the result of a struggle between competing human interests and values (Feenberg 2010). This approach succeeds in explaining why capitalist values are predominantly at work in technologies but is unable to describe how these values become operative in concrete human-technology interactions. I show how a postphenomenological framework (Ihde 1990; Verbeek 2005) is able to complement a neo-Marxist approach in offering a political analysis of concrete technologies. I show how technologies such as social media constitute new normative situations in which new values arise. In doing so, this paper contributes with a phenomenological account of how phenomena such as popularity contests, outrage culture, devaluation of truth, and trolling culture have come to dominate the contemporary public debate online.

References

Bucher, Taina. If...then: algorithmic power and politics. New York: Oxford University Press, 2018.
Feenberg, Andrew. Between reason and experience: essays in technology and modernity. Inside technology. Cambridge, Mass: MIT Press, 2010.
Gertz, Nolen. Nihilism and technology. Lanham: Rowman & Littlefield International, 2018.
Ihde, Don. Technology and the lifeworld: from garden to earth. The Indiana series in the philosophy of technology. Bloomington: Indiana University Press, 1990.
Marres, Noortje. 'Why We Can't Have Our Facts Back'. Engaging Science, Technology, and Society 4 (24 juli 2018): 423. https://doi.org/10.17351/ests2018.188.
Phillips, Whitney. This Is Why We Can't Have Nice Things: Mapping the Relationship between Online Trolling and Mainstream Culture. The Information Society Series. Cambridge, MA; London: The MIT Press, 2015.
Verbeek, Peter-Paul. What Things Do: Philosophical Reflections on Technology, Agency, and Design. Translated by Robert P. Crease. University Park, Pa: Pennsylvania State Univ. Press, 2005.

Session 4F: Using Participatory Value Evaluation to reach unorganised affected stakeholders of military autonomous systems

Christine Boshuijzen-van Burken, Shannon Spruit, Lotte Fillerup & Niek Mouter

The 'value sensitive design' (VSD) methodology, initially developed by Friedman and Kahn in the 1990's, is an iterative three partite design methodology existing of conceptual, empirical and technical investigations (Friedman & Kahn, 2003). In the research here presented, we introduce two novel methods for eliciting values from stakeholders, namely a Group Decision Room (GDR) (used by e.g. Doorn, 2012) and participatory value evaluation (PVE) (Mouter et al., 2019) which allows for engaging unorganized affected stakeholders in a VSD process. Engaging unorganized affected stakeholders is important for the broader research project that this proposal is part of, namely the development of an ethical framework for the design of autonomous systems in Defence.

We refer to Ulrich's distinction between affected and involved stakeholders (2010). Affected stakeholders are people who experience the outcomes of technical (e.g. an artefact), socio-technical (e.g. city planning), or policy design and who are not typically involved (ibid) into a design process, e.g. via the provision of funding, or being the engineer or compliance offer. In some instances, affected stakeholders are formally or informally organized, e.g. in associations, pressure groups, or via social media groups. Unorganized stakeholders lack a spokesperson or identifiable outward facing office or website, which makes them difficult to reach. A GDR allows for structured online anonymous discussion between stakeholders, taking away power dynamics (through known hierarchies, affiliation, extrovert/introvert personalities) that may be present in traditional focus groups. PVE allows for reaching unorganized stakeholders.

The context for our project is Australia. We rely on values that are important to the design of autonomous military systems as expressed by academics, policy makers and pressure groups in autonomous systems debates (see e.g. (Hagendorff, 2020; ICRC, 2021; NATO, 2021; Walsh, 2015)), but in particular the 'silent majority' of Australian citizens. Initial findings from the GDR reveal a general concern for environmental values, geo-political and economic stability. A PVE based on these and other values is designed around an autonomous mine counter underwater vessel and an autonomous drone that drops bombs. Initial results from the PVE are expected to be available by the end of March 2023. Our approach could potentially be used as a blueprint to track changing values in time and between geographically and culturally diverse audiences.

References

Doorn, N. (2012). Exploring Responsibility Rationales in Research and Development (R&D). Science, Technology, & Human Values, 37(3), 180–209. https://doi.org/10.1177/0162243911405344
Friedman, B., & Kahn, P. (2003). Human Values, ehics and design. In The human-computer interaction handbook (pp. 1177–1201). https://brandorn.com/img/writing/tech-ethics/human-values-ethics-and-design.pdf
Hagendorff, T. (2020). The Ethics of AI Ethics: An Evaluation of Guidelines. Minds and Machines, 30(1), 99–120. https://doi.org/10.1007/s11023-020-09517-8
ICRC. (2021). ICRC Position Paper: Artificial intelligence and machine learning in armed conflict: A human-centred approach. International Review of the Red Cross, 913. http://international-review.icrc.org/articles/ai-and-machine-learning-in-armed-conflict-a-human-centred-approach-913
Mouter, N., Koster, P., & Dekker, T. (2019). An Introduction to Participatory Value Evaluation. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3358814
NATO. (2021, October 25). NATO Review—An Artificial Intelligence Strategy for NATO. NATO Review. https://www.nato.int/docu/review/articles/2021/10/25/an-artificial-intelligence-strategy-for-nato/index.html
Ulrich, W., & Reynolds, M. (2010). Critical systems heuristics. In Systems approaches to managing change: A practical guide (pp. 243–292). Springer.
Walsh, T. (2015). Autonomous Weapons: An open letter from AI & robotics researchers. Future of Life Institute. https://futureoflife.org/open-letter-autonomous-weapons/

Session 4F: Why do we need Norm Sensitive Design? A WEIRD critique of value sensitive approaches to design

Diana Adela Martin, Rockwell F. Clancy, Qin Zhu & Gunter Bombaerts

Our contribution argues that mainstream value-sensitive approaches to design have been based on narrow understandings of personhood and social dynamics, which are biased towards Western Educated Industrialized Rich and Democratic (WEIRD) cultures and challenged by empirical evidence. To respond to this weakness, the article suggests that design may benefit from focusing on user behaviours from the joint perspective of values and norms, especially across cultural contexts. For that reason, we propose Norm Sensitive Design as a complement to value-sensitive approaches when designing and implementing new technologies. Versus values, norms serve as more accurate predictors or descriptors of behaviours and can thus support value-sensitive approaches to realize the aspiration of informing user behaviour via design.

Our contribution is structured into three parts: first, it discusses the history of and motivations for Value Sensitive Design and Design for Values; next, it provides an empirical critique based on a non-WEIRD perspective; thirdly, it introduces Norm Sensitive Design as a complement to value-based approaches, presenting mosh pits as an example of an artefact that highlights the need to consider norms alongside values in design. The contribution concludes with a list of reflective questions to aid designers and design researchers in engaging with Norm Sensitive Design, which instructors can also use in the classroom when teaching design ethics.

Our proposal has two major implications. First, on the theoretical side, it makes the case for considering norms alongside values in design, introducing the concept of Norm Sensitive Design. Second, on the practical side, it offers designers and instructors alike prompts for reflecting on design ethics. Overall, our proposal for bringing norm-sensitive approaches to design is a call for making explicit with regard to value-based approaches practices that are already carried out in the fields of behavioural economics and social psychology, or when developing decision architecture and formulating “nudges.”

Norm Sensitive Design can aid value-sensitive approaches to understand and inform collective behaviours, which would be impossible in terms of values alone. Our proposal is thus envisioned to contribute to an enhanced awareness and attention to how we incorporate WEIRD perspectives and practices through ethical lenses in technological designs. This is of heightened importance given that design has been passing along value-laden assumptions and perpetuating societal norms of behaviour, to the exclusion of historically marginalised cultural groups.

Session 5: The Politics—and Philosophy—of Prestige in Engineering

Guru Madhavan

Using examples from the history of mission-critical engineering and the contemporary nature of large-scale programs, this talk will build on how we frequently recognize one act and aspect of engineering (innovation) over another (maintenance) across our careers, corporations, and culture. It will consider some forces that shape the politics—and philosophy—of prestige associated with engineering priorities and projects. In a world pulled by and pushed toward the novel and the nifty, how can commercial innovations gain more grounded vision informed by a sensibility of care and conservation?

Session 5A: AI, Robots and Value Change: Exploring Anthropomorphism’s Normative Impact on Day-to-Day Social Practices

Tom Coggins & Olya Kudina

Nowadays, many AI-powered information communication technologies (ICTs) simulate what it is like to interact with another person. For instance, virtual assistants, social robots, and chatbots often create the impression that they understand and can respond to utterances as a human would. Research shows that people often treat ICTs of this kind as though they are social actors, thus assume that they have psychological attributes they do not possess, such as interests, personalities, or opinions. Researchers usually call this phenomenon “anthropomorphism”. In this workshop, we will examine anthropomorphism via the literature on value change. In the first part of the workshop, we introduce relevant theoretical insights to demonstrate that seemingly socially cognizant ICTs alter how people realize social practices they find valuable. For instance, research shows that virtual assistants installed inside homes influence how occupants communicate to one another and manage tasks including childcare and housework. Likewise, social robots deployed in retail environments may create new workplace expectations associated with emotional labor that human services workers sense that they must meet. We posit that these ICTs (and others like them) encourage their users to adapt pre-existing social practices around them to ensure that they can complete normatively significant tasks. In the second part of the workshop, we will help participants analyze a selection of cases involving anthropomorphism via the value change literature discussed earlier. During this exercise, participants will envision how people would respond to the presence of an ICT that simulates what it is like to interact with another person, discuss the impact said technology will likely have on pre-existing social practices, and evaluate whether such changes are ethically justifiable.

Session 5B: Visions of Change. Battling perfect moral storms in the context of energy transition

Anna Melnyk & Joost Alleblas

In 2006, Stephen M. Gardiner argued that climate change constituted a perfect moral storm. He made an analogy with the true, and tragic story of the Andrea Gail, a vessel at sea facing three converging storms. Gardiner identified global, intergenerational, and theoretical ‘storms’ converging in the ‘ethical tragedy of climate change’. The first two storms are the result of two key dimensions characterizing climate change: its geographical and temporal dimension. Both these dimensions lead to the dispersion of causes and effects, the fragmentation of agency, and institutional inadequacy. Regarding the third storm, Gardiner argued that human theories and knowledge are insufficient to comprehend and deal with climate change effectively.

In this workshop, we focus on a lack of visions of sociotechnical change needed for the energy sector. We perceive this lack as one form of institutional inadequacy that leaves societies incapacitated to adapt to, and proactively deal with, climate change. Visions of sociotechnical change motivate and coordinate different stakeholders. Visions also allow societies to discursively address the future(s). Furthermore, visions that communicate collectively desired futures, enable actors to navigate difficult questions concerning the energy system on critical junctures (and crises) in the future.

Currently, however, on a European political level, these visions seem to be lacking. Decisions are often made in the crisis-mode, leading to ad-hoc solutions to address urgent problems. Examples are the construction of biomass-energy plants in urban areas, intensified coal-mining in Germany and, some would argue, the building of two nuclear reactors in the Netherlands. Despite the absence of a guiding vision of a desirable and feasible future, the above choices have long-term societal consequences. Not only do these technologies bring about certain concerns and values, such as safety, security and health, they might also lead to changing values mid- to long term.

In this workshop, we will experiment with and explore different visions of the future of the energy systems in the mid- to long term (2040). We will use a scenario method to explore different futures of the renewable energy transition. We will particularly focus on how these different scenarios induce potential trade-offs between environmental and human values and impact value prioritization in different energy transition pathways.

Session 5C: Techno-Moral Change and the Role of Scenario Thinking and Modelling in Solar Climate Engineering Research

Benjamin Hofbauer, Behnam Taebi, Giacomo Marangoni & Gideon Futerman

Rapid global warming necessitates radical, swift, and decisive action in terms of both mitigation and adaptation efforts. Yet, the severity of climate impacts (i.e. droughts, floods, rising sea levels, etc.), paired with the socio-political as well as technological challenge of deep decarbonisation, has prompted some scholars and policy makers to consider climate engineering technologies as an addition to mitigation and adaption strategies. In principle, there are two ways of actively engineering the climate. One, by reducing atmospheric carbon; and two, by reducing incoming solar radiation through so-called Solar Radiation Modification (SRM). SRM is presumably capable to deal with an important impact of climate change: i.e. (global) warming.

Throughout the workshop, we will focus on Solar Radiation Modification, specifically through Stratospheric Aerosol Injection (SAI). The SAI is considered a feasible and fairly easy to deploy technology, but research and deployment of SAI are extremely controversial, raising value questions of justice, hubris, and the human/nature relationship, among others. We will explore possible future scenarios of how the research (and potential deployment) of such technologies could shape fundamental societal values, and how those values are reflected in current modelling approaches (e.g. Integrated Assessment Models, IAM). As a theoretical baseline, we will introduce approaches from the techno-moral change and techno-moral scenario building scholarship, and give an outline of how modellers account for values.

The first part of the workshop will be used to introduce the relevant technologies, as well as introduce techno-moral change alongside IAM. Participants will then have time to create their own scenarios, before engaging in a plenary debate and discussing their thoughts. The aim of the workshop is to develop a more holistic understanding of how climate engineering technologies are not mere tools, but manifestations of value-laden and often competing worldviews, and how values play a role when it comes to modelling. What kind of worldviews they could represent will be the guiding question to explore for the participants and organizers.

Thursday

Session 6A: A Value Sensitive Design (VSD) approach for the sustainable design of multiscale CO2 electrochemical conversion (CO2ER

Marula Tsagkari, Mar Pérez-Fortes & Ibo van de Poel

The current study uses a value-sensitive design (VSD) approach to contribute to CO2 Electrochemical Reduction system design (CO2ER), which is within the frame of Carbon Capture and Utilization (CCU) technologies using renewable energy to convert CO2 into valuable products. The focus is on a technology that is at low technology readiness level; by bridging perspectives of stakeholders across the supply chain, and by proactively identifying the relevant values and potential value tensions, the study pinpoints sustainability-related values that are crucial to consider at early design stages. .
The VSD approach has been chosen because it is a methodology that can be applied early in the technology design process, it facilitates communication between engineers and social scientists, and it brings together proactively various considerations of multiple stakeholders. The research uses a two-step VSD approach; 1. identification of relevant norms associated with CO2ER via literature analysis, and 2. qualitative interviews with relevant stakeholders. Then, network analysis is used to identify actors that share common values and actors with different values that may be difficult to collaborate. The current work fills a gap in the research of VSD on low-emission technologies such as CO2ER, while providing valuable information for the successful scaling-up of CO2 electrolyzes.

The results show that the sustainability-related values care for nature, climate change mitigation, use of renewable energy critical materials cost and return on investment, are the most relevant for all the actors, though they may be interpreted differently. For example, values of ownership and avoidance of carbon lock-in were expressed differently among actors, highlighting different expectations and norms. The stakeholders are grouped into 3 clusters according to their main value orientation and the number of values they share with other stakeholders: (i) mostly non-for-profit organizations and academics who represent strong sustainability views and share few values with other actors, (ii) technology implementers, mostly start-ups and industries who share multiple values among them with a strong technological and market orientation and, (iii)an heterogenous group of indirect actors. These results offer useful insights to fulfill the expectations of the various stakeholders in the CO2ER supply chain. More research and continuous engagement with the stakeholders and their emerging or reshaped values will be necessary throughout the design process.

Session 6A: Value Change by Ethical Vision Design: Integrated ethics methods for developing the research agenda in an interdisciplinary engineering project – A field report from the research network PureMobility

Sabine Ammon, Nele Fischer, Tim Hildebrandt, Steffen Müller, Dieter Peitsch & Utz von Wagner

This interdisciplinary contribution brings together different disciplinary backgrounds, namely philosophy, vehicle development, aeronautics, and mechanical engineering. It is a field report from an (ongoing) process of developing a research agenda for an interdisciplinary research network on new approaches in urban and sub-urban mobility. For the participating engineers this project has been a first-time encounter of working closely with a philosopher who has been embedded in the process right from the start. In developing the research agenda, integrated ethics, and more precisely, ethical vision design (Ammon 2020) played a major role for shaping a joint conceptual space based on value change (Van den Poel 2021). In our contribution we will both describe how we proceeded and reflect upon the role integrated ethics can play in such a process.

The aim of the project team was to develop a research agenda which aims at a paradigm shift in mobility. Mobility is central to modern societies. It is an essential part of quality of life and offers a great number of chances and opportunities. It is thus an import factor for prosperity and content. However, despite a multitude of technological developments, traffic still results in an enormous consumption of space, energy, and resources. The omnipresence of – often unused - vehicles and the land consumption of traffic infrastructure cause increasing problems especially in urban and suburban areas. CO2, pollutant emissions, noise and safety are only some of the problems we are facing today with mobility. Present day mobility is a pressing example for the need of value change!

To approach this challenge, the method of ethical vision design is structured in three phases: 1. vision design based on value change, 2. developing conceptual constraints of the design space, 3. outline ethical design challenges in the form of case studies based on issues of justice in current mobility patterns. Early in the process, the team developed a research vision based on value change in transport systems which helped to structure the following steps. In interdisciplinary deliberations we extracted core values which stand for a fundamental value change in transport systems. Mobility should only use minimal space, energy and resources to meet the mobility needs of our society. This allowed us to define “PureMobility” by values like human-centered, clean, essential, safe, sustainable, and efficient. However, making this value-based vision operative for the research agenda proves to be difficult. Only by introducing conceptual constraints of the design space allows to proceed in a meaningful way. Based on this, we are able to delineate ethical design challenges which are based on condensed case studies to explore ethical issues throughout the development process.

Ethical vision design allows to reframe the design problem; instead of technology push based on technological feasibility we observe a vision pull. By sharing both our learnings and methodological approaches, we want to discuss how ethical vision design based on value change offers an approach to overcome engineering approaches which are exclusively grounded in technological solutionism.

References

Ammon, Sabine 2020. Ethical Vision Design im BERLIN ETHICS LAB: Technologievisionen in der Entwicklung verantwortlicher KI und verantwortlicher Mensch-Maschine-Interaktion. In: Interdisziplinäre Arbeitsgruppe Verantwortung: Maschinelles Lernen und Künstliche Intelligenz der Berlin-Brandenburgischen Akademie der Wissenschaften (ed.): KI als Laboratorium? Ethik als Aufgabe!, Berlin Brandenburgische Akademie der Wissenschaften, pp.10–14.https://www.bbaw.de/files-bbaw/user_upload/publikationen/BBAW_Verantwortung-KI-3-2020_PDF-A-1b.pdf
Van de Poel, Ibo 2020. Values and Design. In: Neelke Doorn, Diane Michelfelder (ed.): Routledge Hand-book of Philosophy of Engineering, pp. 300–314.

Session 6A: Safe-by-Design: The need for engineered prerequisites for inherent safety

Britte Bouchaut, Thom Tribble & Lotte Asveld

Fighting pollution by Per- and polyfluoroalkyl substances (PFAS) and developing safe alternatives to this group of chemicals requires new types of chemistry. Biotechnology and synthetic biology can contribute to finding such. However, like with every new technology or application, new risks might emerge and we must ensure that these new alternatives are no less harmful than those they replace (i.e. regrettable substitution) (Bouchaut et al., 2022; Zimmerman & Anastas, 2015). In literature, the Safe-by-Design (SbD) approach is argued to be a way to prevent such from happening as it provides guidelines and an iterative approach to identify possible emerging risks and develop anticipatory strategies to lower or mitigate these, during early stages of development (Bouchaut & Asveld, 2021; Robaey, 2018). Thereby, SbD relies heavily on knowledge of properties and characteristics of materials and chemicals to be able to identify potential emerging risks and to define appropriate risk-lowering strategies – thereby creating inherent safety.

In this paper, we have researched the monitoring and regulation of PFAS in the Netherlands and Europe. It is known now that PFAS degrade extremely slowly due to their stability and persistence in the environment, and that these chemicals accumulate in humans, resulting in adverse health effects (Schrenk et al., 2020). This evidence is strongest for long-chain PFAS, and most are restricted under the Stockholm Convention (UN Environment Programme, n.d.). In response, industry has developed and produced short-chain PFAS as a safer alternative. However, new studies are showing that these too have adverse effects on human health, but this is difficult to measure as there appears to be no suitable analytical equipment to do so.

This illustrates a prerequisite for successfully implementing SbD: we need to have equipment (e.g. analytical tools for monitoring and measuring) that generates knowledge needed for SbD to be successfully incorporated. Without such, potential risks cannot be estimated during early stages of development, nor can anticipatory strategies be developed which is crucial for new types of (synthetic) chemicals and biotechnologies to be developed safely and responsibly.

References

Bouchaut, B., & Asveld, L. (2021). Responsible Learning About Risks Arising from Emerging Biotechnologies. Science and Engineering Ethics, 27(2), 22. https://doi.org/10.1007/s11948-021-00300-1
Bouchaut, B., Hollmann, F., & Asveld, L. (2022). Differences in barriers for controlled learning about safety between biotechnology and chemistry. Nature Communications, 13(1), 1–4. https://doi.org/10.1038/s41467-022-31870-8
Robaey, Z. (2018). Dealing with risks of biotechnology : understanding the potential of Safe-by-Design [Report commissioned by the Dutch Ministry of Infrastructure and Water Management, The Hague, The Netherlands]. https://doi.org/10.13140/RG.2.2.13725.97769
Schrenk, D., Bignami, M., Bodin, L., Chipman, J. K., del Mazo, J., Grasl-Kraupp, B., Hogstrand, C., Hoogenboom, L., Leblanc, J. C., Nebbia, C. S., Nielsen, E., Ntzani, E., Petersen, A., Sand, S., Vleminckx, C., Wallace, H., Barregård, L., Ceccatelli, S., Cravedi, J. P., … Schwerdtle, T. (2020). Risk to human health related to the presence of perfluoroalkyl substances in food. EFSA Journal, 18(9), e06223. https://doi.org/10.2903/J.EFSA.2020.6223
UN Environment Programme. (n.d.). Stockholm Convention: Overview PFAS. Stockholm Convention. Via http://chm.pops.int/Implementation/IndustrialPOPs/PFAS/Overview/tabid/5221/Default.aspx
Zimmerman, J. B., & Anastas, P. T. (2015). Toward substitution with no regrets. Science, 347(6227), 1198–1199. https://doi.org/10.1126/science.aaa0812

Session 6B: Confronting the shadows of colonialism in engineering education: An ongoing practice of “Calling In”

Dimpho Radebe & Kai Zhuang

The murder of George Floyd sparked increased global attention on the Black Lives Matter movement, and ongoing demands to increase diversity and create more inclusive environments within engineering spaces. It also illuminated the ways in which the “past-present histories” of slavery and colonialism continue to manifest through neoliberalism and racial capitalism today. However, this awareness, for many, was brief before other priorities re-emerged. Thus, the authors recognise that the current approaches to increasing awareness and consciousness of the legacy and traces of colonial harm globally, and within the field of engineering specifically, are insufficient.

The fields of psychology and neuroscience have identified that, while individual experiences may vary, this harm has manifested as collective intergenerational and intersectional trauma. We also recognise that societies intentionally and unintentionally choose to keep this trauma hidden from our collective conscious view and decision-making. In other words, it remains hidden in our collective “shadow”. Looking at one’s shadow requires engaging with often painful emotions (e.g., fear, guilt, shame) that are difficult to process. It often has us questioning what we thought we knew and why we knew it – and that uncertainty is a very uncomfortable place to be in. It requires engaging with a part of us that very often comes into being as a protection mechanism due to a past traumatic experience. Consequently, by engaging with that part of ourselves, we often fall into survival mode: choosing fight, flight, freezing, or fawning as a response. The sensitive, emotional nature of these discussions combined with the painful feelings that arise through this work, make it hard to do. Thus, as engineering educators, if we truly want to help ourselves and our students confront these “past-present histories” that continue to manifest within our field today, we need a framework in which to navigate and explore these issues.

In this conceptual paper, we present a framework to help engineering educators navigate through these conversations. We first use this framework to provide examples of how the impacts of colonialism continue to manifest and show up in engineering. We explore the ways in which dominant narratives such as the socio-technical dualism and engineering mindsets embedded within engineering education form dominant paradigms that inadequately prepare engineers and engineering educators to engage effectively with our collective shadow. We then use this framework to reflect separately, using an autoethnographic method, on our own individual experiences in engineering to demonstrate its use. Finally, we posit that for engineering as a field to engage with our shadows in ways that the fields of health, law, and education are doing so as well, we need to “call in” the field of engineering collectively to create a “moral infrastructure”. That is, we need to be willing to engage emotionally with the ways in which our field is contributing to inequities and human and environmental exploitation to create positive transformational change. We conclude with some examples of how we are currently implementing this framework through workshops and retreats to support engineering educators and students.

Session 6B: Rethinking McMindfulness from an East Asian Perspective

Joseph Emmanuel D. Sta. Maria & Matthew J. Dennis

Mindfulness apps are increasingly popular in the self-care app market. Since 2018, when Apple announced that self-care apps were the App Trend of the Year, the popularity of these products has increased exponentially (Apple 2018). Nevertheless, mindfulness apps have been accused of “McMindfulness,” a catch-all term signifying their propensity to make users subservient to the socioeconomic status quo (Hyland, 2017, Purser, 2019, Slunecko and Clouba, 2021). Critics of McMindfulness contend that mindfulness apps merely serve as palliatives that help people cope with unjust structures, instead of aiding them in evaluating and critiquing these structures. This is because the meditative practices promoted by McMindfulness apps are devoid of any ethical orientation (Cannon, 2016, Simão, 2019). In contrast, the non-Western traditions from which mindfulness originated understood meditative practices as having an ethical purpose (Garfield, 2017, Peng and Zhang, 2022). There have thus been calls from the critics of McMindfulness for a more socio-ethically aware version of mindfulness (Stanley, 2012, Magee, 2016, Walsh, 2016, Leggett, 2022). For some scholars, this can be achieved if mindfulness practices return to their originary, Buddhist, ethical principles (Purser and Loy, 2013, Hyland, 2017). However, others argue that, even from the Buddhist perspective, mindfulness practices are not meant to evaluate socio-ethical issues or structures (Repetti, 2016, Analayo, 2020). This paper aims to respond to the call of McMindfulness critics. While the Buddhist tradition is useful, the paper argues that Confucian philosophy is better equipped for formulating a socio-ethically aware and critical mindfulness practice. This is because Confucianism’s conception of mindfulness is more explicitly oriented toward socio-ethical matters, and it also encourages critical investigation (Tiwald, 2018, Tan 2019, 2021). The paper explores how Confucian philosophy can offer insights for improving the design of the next generation of mindfulness apps, such that they avoid McMindfulness and, instead, actuate mindfulness's socio-ethical potential.

Session 6B: The “waiting room version of history” in post-colonial- and sustainable-engineering projects

Andrea Gammon

What conceptions of time and history are at work in engineering projects? This paper takes its inspiration from historian, postcolonial theorist, and subaltern studies scholar Dipesh Chakarabarty and investigates the ways in which Western, especially European, conceptions of history pervade engineering works in the past and present. I defend how, similar to values, temporal and historical orientations in engineering projects can be understood as expressing larger, structural, and often global ideas and forces. Large-scale engineering works in the post-colonial period employed a straightforward, if problematic, temporal conception. Motivated by ideas of modernization and progress, these projects promised development: a future of modern ways of life and improved health, opportunity, and prosperity. I argue that these projects, as development projects, exemplify what Dipesh Chakrabarty terms the waiting room version of history: that is, they aim at closing the cultural distance between the “developed” Global North, and the “underdeveloped” Global South, a distance based in European conceptions of history and progress. I then look to more contemporary engineering projects, where environmental sustainability takes precedence, to ask if the temporal orientation in older, post-colonial era projects is challenged as development is recast as sustainable development and environmental sustainability replaces modernization as a main objective. I argue that the waiting room version of history continues to shape how expertise is constructed: what factors are taken as relevant for engineering knowledge, and who shares and receives knowledge in sustainability-driven engineering projects. I also motivate the focus on such engineering projects as interventions into the environment that influence understandings of and encounters with nonhuman nature.

Works Cited

Chakrabarty, D. 2000. Provincializing Europe. Princeton University Press.

Session 6C: The ethics of engineering vs. the politics of technology

Maarten Franssen

Value questions have always been central to the philosophy of technology. However, the field lacks a well-founded view on how to tackle them. In the past 25 years, 'ethics of technology' has become the single label under which normative issues are studied. This is highly problematic, because the perspective of ethics is of an individual acting in the world. Ethical theories concern whether an individual agent's actions are right or wrong, 'to do' or 'to do not' given some duty, and which values an individual agent's action will honour or harm. Certainly this perspective can be brought to bear on technology, in the form of professional ethics – the ethics of engineering. Many normative issues related to technology, however, concern the development and implementation of technologies, their societal impact and their regulation and control, all of which concern the societal, that is, aggregate level. There we lose moral track of the actions of individuals. Even within engineering ethics it is already problematic to what extent the design of technologies can be seen as under any individual's control, but although this is occasionally acknowledged (e.g. Van de Poel & Royakkers 2011), still exclusively the conceptual framework of ethics is offered for structuring the discussion.

In this paper, I make a plea for introducing a radically different approach to many normative and value-related issues raised by engineering and technology: through political philosophy rather than ethics. By political philosophy I mean approaches that target the normative structure of society, the rules, rights and duties that define the arena for the actions of its individual members. Philosophically this work is currently dominated by the contractarian approach introduced by Rawls (1971/1999). In arguing the relevance of this approach, I point out its shortcomings for dealing with technology as well. Rawlsian moral-political philosophy identifies the individual agents making up society as the only entities that act, and 'society' as their mere collection, and argues for a principle of maximal liberty for agents. What needs to be developed in addition is (1) a conception of business firms as supra-individual corporate agents, since virtually all technology is developed by such agents, plus (2) a conception of the state as agent, since the state also plays a crucial role in technology development and control (Mazzucato 2013), plus (3) a conception of the sovereign state as a member of a global community of sovereign states, since many of the great issues of today (e.g. global warming, geo-engineering) can be meaningfully addressed only at that level. Rawls's attempt (1999) to cover the latter aspect is inadequate. Finally it is important to address the challenge that technology poses to the contractarian approach itself, since it "rests on an assumption of equal rationality among persons which differences in technology deny" (Gauthier 1986).

References

Gauthier (1986), Morals by agreement, Oxford UP.
Mazzucato (2013), The entrepreneurial state, Anthem Press.
Rawls (1971/1999), A theory of justice, Harvard.
Rawls (1999), The law of peoples, Harvard.
Van de Poel & Royakkers, (2011) Ethics, technology, and engineering: an introduction, Wiley-Blackwell.

Session 6C: How democratic debates contribute to a better governance of innovation

Jérémie Supiot & Philippe Estival

Ensuring an adequate technological development in our modern societies is a particularly "wicked problem" (Pohl et al. 2017). In fact, societal and ecological stakes are often tremendous. They are considered as issues demanding the consideration of extremely various types of information, most of the time highly technical. In this contribution we draw on the limits of technocratic approaches (JOULIN, 2022) of such wicked problems to explore the challenge of technical democracy (CALLON, LASCOUMES, BARTHES, 2001), the need of ethical capacity building of citizens, and the key role that could play online debate platforms .

Although it is the responsibility of engineers to find technical solutions, the ethical evaluation of environmental and social impacts should be a political question for all citizens. In this regard, ethics by design is insufficient (Fischer, 2019) because it does not involve democratic consultations. However, many citizens do not have the technical skills to understand fully the limits and impacts of technologies, and it is impossible to ask every citizen to evaluate all new technologies. Moreover, most new technologies are developed at the discretion of private companies, and are regulated only after they are put on the market (Bonneuil, Fressoz, 2014). How to get democratic control of innovation, regarding their social and environmental impacts ?

At a time when social dumping and law shopping make it impossible to regulate efficiently at the international level (for example, many neonicotinoids are still on the market despite their devastator impacts on pollinators), Corporate Social Responsibility seems to be the only way for citizens to influence the development of new technologies. However, most of the time, the only role left to citizens is boycott or investment, relying on the social and environmental impacts companies accept to share publicly, hoping that it is not just greenwashing (Berlan, Carbou, Teulières, 2022) . Moreover, tremendous investments are made in technologies without involving any stakeholder in the process, leading to path dependencies (David, 1980) for unneeded technologies(Bonneuil, Fressoz, 2014) like metaverses for instance. This situation can be analyzed as the result of the technocratic ideal (Joulin, 2022) that lead to the "governance by numbers" (Supiot, 2020) that we see in many companies and countries. Because regulation is always more difficult after they are put the market, we should find a democratic solution for regulating technologies at the very beginning : before they are designed.

We will argue that governance of innovation is not about social acceptability, but about democratic debates like hybrid forums (Callon and al., 2009). Those forums can be held on digital platforms (Seilles and Al., 2011) like Cartodébat, allowing all citizens to contribute to debates on controversial technologies. We will argue that online debate platforms can be designed to improve the quality of argumentation allowing participants to build a nuanced viewpoint on new technologies (Seilles & Al., 2011). However, we will also present some of the issues that need a philosophical examination, concerning the legitimacy of the arguments that are published on the platform.

Session 6C: Perfectionism: Towards a political philosophy of technology and human flourishing

Avigail Ferdman

The discourse on ethics of technology tends to focus on preventing harms. Yet there is growing interest in an ethics of technology that explicitly promotes the good: the social good (e.g. justice, freedom) [1] or human flourishing [2].

This paper proposes to take the discussion a step further by offering a political philosophy of technology and the good. It examines the moral requirements from technology, and the corresponding social and political basic structures necessary to ensure human flourishing.

The paper offers a normative account that is grounded in a 'perfectionist basic structure of society'. Perfectionism, in philosophy, holds that our well-being is determined by how well we develop and exercise our human capacities (e.g. to know, be sociable, exercise willpower). The basic structure of society is the network of institutions that forms the background within which individuals and associations interact with one another [3]. Because the basic structure of society is highly involuntary, a moral requirement of justice is that it ought to create the conditions for developing and exercising human capacities [4].

From these premises, the paper argues first that technology is a feature of the basic structure, since it creates affordances and constraints [5] that encourage or limit the development and exercise of human capacities. Second, that technology should create the conditions for developing and exercising human capacities. Third, that there are actually two versions of the perfectionist basic structure. A weak version holds that the basic structure should not create the conditions that constrain the development and exercise of human capacities (e.g. should sanction addictive-by-design technology). A strong version holds that the basic structure should create the conditions that encourage persons to develop and exercise their capacities (e.g. should prioritize open-code systems).

Each version has its strengths and drawbacks. The strong version might be criticized as paternalistic, but its strength is that it is a framework for actually realizing flourishing rather than merely enabling its potential. The weak version is less paternalistic, but it focuses on reducing harms rather than promoting the good. While both versions provide an ethical framework for evaluating technological developments, the strong version is bolder in its vision for technology as a contributor to human flourishing.

References
[1] Philip Brey, “The Strategic Role of Technology in a Good Society,” Technology and the Good Society 52 (2018): 39–45; Luciano Floridi et al., “How to Design AI for Social Good: Seven Essential Factors,” Science and Engineering Ethics 26, no. 3 (2020): 1771–96.
[2] For example Brey, “The Strategic Role of Technology in a Good Society”; Bernd Carsten Stahl, “Concepts of Ethics and Their Application to AI,” Artificial Intelligence for a Better Future: An Ecosystem Perspective on the Ethics of AI and Emerging Digital Technologies, 2021, 19–33; Mirko Farina et al., “AI and Society: A Virtue Ethics Approach,” AI & SOCIETY, 2022.
[3] John Rawls, Political Liberalism (New York: Columbia University Press, 1996).
[4] Avigail Ferdman, “A Perfectionist Basic Structure,” Philosophy & Social Criticism 45, no. 7 (2019): 862–82.
[5] Brey, “The Strategic Role of Technology in a Good Society.”

Session 6D: Datafied doubt as an epistemic virtue: learning from activists’ and practitioners’ perspectives

Dmitry Muravyov

With the proliferation of (big) data, scholars and the public have addressed different types of data-related concerns. On the one hand, many have discussed what ethical approaches may be suitable for working with data, especially when the data is about humans. When people are rendered as data, the quantification and abstraction, associated with it, may become insensitive to humans as subjects and thus be ethically problematic (Thylstrup, 2022). On the other hand, other scholars have addressed data-related epistemological issues by conceptualizing how data, humans, and reality may be related. Scholars have put forward alternative vocabularies to describe this relationship when criticizing the naively positivist view of data as a self-evident representation of an outside world (van Dijck, 2014).

In this paper, I further bridge these ethical and epistemological concerns by proposing the epistemic virtue of datafied doubt. I contend that epistemic virtue as a notion helps to weave together questions of knowledge, subjectivity, and normativity in a way that helps to elicit ethically sensitive questions in working with data and guide collective reflection on them. To illuminate the role of datafied doubt, I propose seeing data not as an object, but as a relation in three senses: recognition, composition, and production.

This paper contributes to the ethics of data, dynamics of value change as well as the relationship between virtues and technology (Vallor, 2016, Boennik & Kudina, 2020, van de Poel & Kudina, 2022, Boenig-Liptsin, 2022). Specifically, I bring forward a relational understanding of data and synthesize some of the insights from the empirical studies to propose normative arguments for a specific ethical and epistemological transformation of how people can relate to data. In doing so, I approach the emergence of ethical and epistemological concerns in a bottom-up manner (Shklovski & Némethy, 2022).

Rather than being a result of purely conceptual thinking, datafied doubt as an epistemic virtue is rooted in studies of data activists’ and practitioners’ practices and is intentionally positioned against dataism as the dominant paradigm. Datafied doubt takes as a point of departure the impossibility of ethical relation between us and others outside of the emergence of a subject that data offers (Amoore, 2019). By synthesizing empirical research on data activists and practitioners, I define datafied doubt as an epistemic virtue with a disposition towards data that is hopefully pessimistic, open-ended, and rooted in practice.

I further link these three dimensions of datafied doubt to an understanding of data as a relation to demonstrate how datafied doubt can be beneficial in explicating previously less salient dimensions of living with data. The practice-based aspect of datafied doubt brings attention to relations that underlie the production of data. The open-ended nature of datafied doubt helps account for an always-in-principal ambivalent inclusion of entities into the datasets and therefore accentuates an implicit presence of human subjects (Metcalf & Crawford, 2016). Being a hopefully pessimistic disposition towards data, datafied doubt highlights the gap between data and the world, pointing out how the former is a partial achievement.

References
Amoore, L. (2019). Doubt and the Algorithm: On the Partial Accounts of Machine Learning. Theory, Culture & Society, 36(6), 147–169. https://doi.org/10.1177/0263276419851846
Boenig-Liptsin, M. (2022). Aiming at the good life in the datafied world: A co-productionist framework of ethics. Big Data & Society, 9(2), 205395172211397. https://doi.org/10.1177/20539517221139782
Boenink, M., & Kudina, O. (2020). Values in responsible research and innovation: From entities to practices. Journal of Responsible Innovation, 7(3), 450–470. https://doi.org/10.1080/23299460.2020.1806451
Metcalf, J., & Crawford, K. (2016). Where are human subjects in Big Data research? The emerging ethics divide. Big Data & Society, 3(1), 205395171665021. https://doi.org/10.1177/2053951716650211
Shklovski, I., & Némethy, C. (2022). Nodes of certainty and spaces for doubt in AI ethics for engineers. Information, Communication & Society, 1–17. https://doi.org/10.1080/1369118X.2021.2014547
Thylstrup, N. B. (2022). The ethics and politics of data sets in the age of machine learning: Deleting traces and encountering remains. Media, Culture & Society, 44(4), 655–671. https://doi.org/10.1177/01634437211060226
Vallor, S. (2016). Technology and the virtues: A philosophical guide to a future worth wanting. Oxford University Press.
van de Poel, I., & Kudina, O. (2022). Understanding Technology-Induced Value Change: A Pragmatist Proposal. Philosophy & Technology, 35(2), 40. https://doi.org/10.1007/s13347-022-00520-8
Van Dijck, J. (2014). Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. Surveillance & Society, 12(2), 197–208. https://doi.org/10.24908/ss.v12i2.4776

Session 6D: Mind-body-data: Epistemic shift of digital psychiatry

Ekaterina Bogdanova

The pandemic had invigorated the aspirations of a digital mental health market. From mood trackers to clinical systems, new start-ups and researchers are striving to develop technologies that would allow diagnosing, predicting, and monitoring that is more accurate, objective, and timely. One of the emergent propositions, digital phenotyping, promises to deliver those results through “non-invasive” methods of real-time collection and analysis of an individual digital footprint. Data collected from multiple sensors and interaction data – such as the number of calls and texts, browser information, GPS information, gyroscope, and speech emotion recognition – becomes evidence for the symptoms of mental distress and health conditions, whether through causation or correlation. Past the apparent issues of privacy and surveillance (both medical and digital), other philosophical concerns come to the surface – those of new identities, subjectivities, and medical evidence. In this paper, I evaluate how the tools of digital psychiatry mediate the experience of body and mind, health and illness. In particular, how do these technologies drive epistemic changes at both individual and institutional levels? These transformations, furthermore, require a different approach to the ethical framework for the design of digital mental health technologies – the one that would be able to align bio-, psychiatric, and data ethics. I address this problem in the conclusion of this paper.

Session 6D: Ethical-epistemic values driving decisions under uncertainty: The case of hydropower in the United States

Caitlin Grady & Lauren Dennis

Critical infrastructure impacts everyday life in support of providing basic goods and services. One such infrastructure, dams, takes many forms and serves purposes from electricity production to flood mitigation to recreation. Under an uncertain future due to climate change, dams are becoming increasingly challenging to modernize, manage, and predict. Dam management is further complicated by complex interactions across technology, the natural environment, and society.

This work studies the emerging ethical and epistemic landscape of hydropower dam management under an uncertain climate future. Using a coupled-epistemic-ethical analysis, we provide evidence that multiple values shape the approaches in science, technology, and policy for dam management in the United States and that these values differ under conditions of location, dam capacity, and ownership structures. This work leverages interviews of hydropower owners and operators conducted in 2022 that sought to understand decision-making under climate uncertainty. Dams in general, including those that generate hydropower, are often a contentious issue because they raise challenges across a variety of competing values and needs. There are clear ethical implications for the value of water and the value of place. Instead of focusing on what is a long history of ethical inquiry surrounding the premise of whether or not to dam a river, this work strives to bring new perspectives on applied ethics relating to the management of our current infrastructure and how are decisions made in the future. We also highlight new moral considerations for discussion relating to whether or not hydropower has a role to play in the United States as we move to a decarbonized electricity grid.

Highlighting the practitioners' reflections gained through interviews, this work seeks to bridge the divide between academic philosophical discussions on the ethical consideration of dams and dam management in the real world. We will conclude this presentation with a discussion of future research directions and how value transparency and policy changes might seek to mitigate some of the challenging management landscapes for hydroelectricity in the United States.

Session 6E: Reprogenetic technologies, future value change, and the axiological underpinnings of reproductive choice

Jon Rueda

The philosophical exploration of how the design, implementation and adoption of technologies may lead to value change is gaining strong traction in the debates on the ethics of technology (van de Poel, 2021; Vetter, 2021; van de Poel & Kudina 2022; van de Poel, 2022). One particular issue that is receiving increasing attention is how technological advances may induce axiological variations in the future and how to anticipate them (Danaher 2021; Hopster 2022; de Wildt et al., 2022). However, this subject has been largely neglected in relation to emerging reprogenetic technologies, with a few rare exception (Rueda et al. 2022).

In this contribution, I analyse philosophically and systematically various questions concerning the future value changes that can be induced by reproductive and genetic technologies. I will start clarifying the very notions of 'value', 'moral value' and 'value change', and introducing the most important emerging reprogenetic innovations. After that, I shall offer a broad account of the different types of value changes that may be elicited by reprogenetics in the future. Then, to narrow my argument, I will ground the phenomenon of future value change in the case of reproductive choice. Indeed, the prospective axiological fluctuations made possible by reprogenetics are more apparent after attending the co-productive relationship between technologies and choice—as far as technologies create affordances and transform action spaces for decision-makers. Finally, I argue that the best way to anticipate future axiological variation is to reinforce our knowledge of the most important mechanisms through which techno-value change operates.

References

Danaher, J. (2021). Axiological futurism: The systematic study of the future of values. Futures, 132(June), 102780. https://doi.org/10.1016/j.futures.2021.102780
de Wildt, T. E., van de Poel, I. R., & Chappin, E. J. (2022). Tracing Long-term Value Change in (Energy) Technologies: Opportunities of Probabilistic Topic Models Using Large Data Sets. Science, Technology, & Human Values, 47(3), 429-458.
Hopster, J. K. G., 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, porkand ploughs: the structure of technomoral revolutions. Inquiry (United Kingdom), 1–33. https://doi.org/10.1080/0020174X.2022.2090434
Hopster, J. (2022). Future value change: Identifying realistic possibilities and risks. Prometheus.
Rueda, J., Pugh, J., & Savulescu, J. (2022). The morally disruptive future of reprogenetic enhancement technologies. Trends in Biotechnology.
van de Poel, I. (2021). Design for value change. Ethics and Information Technology, 23(1), 27–31. https://doi.org/10.1007/s10676-018-9461-9
van de Poel, I., & Kudina, O. (2022). Understanding Technology-Induced Value Change: a Pragmatist Proposal. Philosophy and Technology, 35(2), 1–25. https://doi.org/10.1007/s13347-022-00520-8
van de Poel, I. (2022). Understanding value change. Prometheus, 38(1). https://doi.org/10.13169/prometheus.38.1.0007
Vetter, P. (2021). Changing Values, Changing Technologies. TATuP-Zeitschrift für Technikfolgenabschätzung in Theorie und Praxis, 30(3), 80-81.

Session 6E: How to represent human values in the Metaverse

Rafael Coimbra & Édison Renato Pereira da Silva

The metaverse promises to be a new connection space, mediating social relations through digital devices. Using different types of realities (AR/VR/XR), human representations will be created, replicated, or mixed in analog-digital environments. A fundamental question throughout the construction of the metaverse is: who will build it and based on what values will the human representation in the metaverse be created?

Currently, there are a few metaverse developers, both in terms of applications and infrastructure, most of which are large technology companies concentrated in a few places in the world. It is possible that the direction of the metaverse is being guided by ethnocentric visions, based on the intrinsic values of the societies in which developers find themselves. Also, part of the environment will likely be constructed by digital tools such as generative artificial intelligence.

Both representations of public and private spaces will depend on actions performed during the construction of the metaverse, enabling the creation of different scenarios.

These scenarios vary according to the degree of insertion or suppression of real or digital layers and may keep equal, distort or amplify currently existing human values. We propose to think of four possible situations, from reality to virtualization, and towards utopia or dystopia, as shown in figure 1: i) equal and inclusive representations of the physical world; ii) global physical problems solved using digital layers; iii) partial or biased representations of the physical world; iv) new asymmetries and exclusionary values through digital tools.

i) In this scenario, real local values would be discussed collectively and democratically and then be faithfully transferred to the virtual world. Because it is utopianly directed, real values considered negative would tend to be suppressed by digital tools.
ii) In this scenario, new virtual layers are added to expand the limits of the real universe, creating new human values. By being utopianly directed, it would be possible to achieve a perfect balance of virtual values.
iii) In this scenario, real values would be transferred in a prejudiced and ethnocentric way to the virtual world, determined by small groups that own technologies. Because it is dystopically directed, real values considered positive would tend to be suppressed by digital tools.
iv) In this scenario, new virtual layers are added to expand the limits of the real universe, creating new human values. Because it is dystopically directed, it is possible that, in the extreme, there is a great concentration of powers, toward a process of virtual domination and colonization.

The decisions taken in the present will be decisive for the representation of human values in the metaverse, so they need to be debated by society. In this article, we will present the details and reflections of each of the scenarios put forth.

Session 6E: Genetic enhancement in bioethical discussions: a computational approach to value change

Tomasz Żuradzki

The discovery of how to alter a sequence of genetic material is one of the most important scientific breakthroughs of the second half of the 20th century. Although the direct manipulation of the genome of living organisms (usually plants for agriculture) was embraced by scientists decades ago, the development of CRISPR/Cas9 method in 2012 is considered a revolution due to its efficiency and cost-effectiveness. In 2015, CRISPR/Cas9 germline modifications were first used in non-viable human embryos, opening a real possibility of making permanent, heritable changes to the human genome.

This paper aims to systematically analyze scholarly discussions in bioethics on genetic enhancement, in particular on human genetic enhancement, and more generally, on ethical and regulatory issues concerning the modification of the genome. A standard manner in which practitioners of an academic discipline reflect upon the development of their field is through "close reading" of selected texts, which is often mediated by their personal experience and academic interests. Here is a typical statement based on such an approach: "enhancement is coming to the forefront of bioethical scholarship" since this topic "combines cutting-edge science with mainstream philosophy" (Harris 2012).

The approach we adopt in this paper takes seriously the epistemological question of how one can justify this type of statement. Referring to our previous studies based on the corpus of about 20.000 texts published since 1971 in seven leading journals in the field of bioethics (Bystranowski, Dranseika, Żuradzki 2022a, b), we use a "distant reading" approach based on topic modeling (a computational text-mining technique aimed at discovering hidden thematic compositions in large text corpora) and citation analysis. In this paper, we concentrate on the topic we previously interpreted as Enhancement (characterized by terms "enhancement," "enhance," "technology," "intervention," "cognitive," "capacity," "trait," "morally," "improve," "bioenhancement"), which was "the biggest winner" in terms of relative growth in our corpus (the increase of the mean prominence from 0.03% in 1971-75 to 0.97% in 2016-20). We also include in our analyses four strictly correlated topics (in the sense of being the most frequently present together in the same texts with the topic Enhancement), that is, Germline, Ecology, Offspring, and Genetics and we delineate a sub-corpus of papers that 'belong' to this five-topics cluster, which we interpret as the core of bioethical discussions on genetic enhancement.

We investigate changes in conceptualizations of values related to genetic enhancement (more on value changes see: van de Poel 2021). In particular, we investigate what is treated in bioethics as the most important ethical or regulatory challenges stemming from scientific breakthroughs concerning genetic enhancement. Our analysis shows that a significant part of discussions on genetic enhancement in bioethics is concentrated on the problems with moral or cognitive human enhancement. Surprisingly, our citation analyses of the sub-corpus show that bioethical discussions on genetic enhancement rarely follow recent scientific breakthroughs and rarely engage with the main philosophical problems that could be important in this context. Therefore, our study may be interpreted as a step to undermine the claim that bioethical discussions on genetic enhancement "combine cutting-edge science with mainstream philosophy".

Session 7A: An UnWorkshop for the Education of a Whole New Engineer

David E. Goldberg, John R. Donald, Beata Francis &
Katherine A. Goodman

In June 2017, an unusual event called the Educational Transformers Unconference (2017) was held at Lehigh University in which a combination of planned and spontaneous activities was facilitated to encourage fresh approaches to the intellectual, emotional, and cultural change and enhancement of university education. A key element of the ETU event was the emphasis on a key insight of Goldberg and Somerville's A Whole New Engineer (2014) stressing the importance of culture and emotion in effective educational change. As do others (Duke Education, n.d.), we use the prefix un- to indicate a shift from conventional academic activities, which are proscribed and intended to disseminate information from an expert to an audience. The un- version of a conference, workshop, or poster session, in contrast, is one that provides a facilitated experience for participants to explore and create knowledge through collaborative shifts in perspective.

Building on the success of the 2017 event, this abstract describes a two-hour unworkshop in which a similar combination of both planned and spontaneous activities is facilitated at the April 2023, Forum on Philosophy, Engineering, and Technology (www.fpet2023.org). The unworkshop also builds on the independent exploration and application of key ideas and practices by four engineering educators with combined experiences across Africa, Asia, Australia, Europe, as well as North and South America.

Squeezing the success of a two-and-a-half-day event into a two-hour workshop is a significant challenge, but the emotional-cultural flow of the original event (Educational Transformers UnConference, 2017) is preserved. The discussion is focused on the education of engineers and technologists and emphasizes key advances in understanding core educational change processes.

The unworkshop is divided into four activity periods of roughly equal duration:
1. Period 1: A brief micro-didactic introduction to co-contraries: design for changing values, followed by interactive pairwork to unearth co-contraries worthy of further exploration in view of identifying alignment of deeper assumptions.
2. Period 2: Planned and spontaneous activities to focus on key co-contraries for the improvement of engineering education through storytelling and conversations.
3. Period 3: An unposter session to tell true or futuristic stories that are fantastic or disastrous for engineering education.
4. Period 4: A closing that emphasizes participant takeaways, injected beauty, and the emotional-cultural shifts of the unworkshop.

The first period starts with a very brief discussion (and short handout) of co-contraries (sometimes called polarities)—or opposites that need each other—using the teaching sequence of Goldberg and Somerville (in press). This is followed by pairwise discussion or pairwork (Goldberg, 2009) of key co-contraries of educational change. Participant insights from this session become the seeds for the second period with several spontaneous facilitated discussions or activities exploring key co-contraries as well as one or two prepared activities involving key co-contraries (freedom && structure, knowing && not knowing, listening && speaking, etc.).

The third period is a short unposter session in which small teams of participants prepare and present a spontaneous unposter that tells a story important to the future of engineering education, a future connected to something participants care about. Materials (flip charts, markers, magazine pictures, scissors, gluesticks, etc) for the creation of the unposter and detailed instructions (a brief handout) for its rapid construction are given to participants to aid this activity.

The final session (fourth period) concludes with participant takeaways from the unworkshop, emphasizing the emotional, cultural, and philosophical shifts that may or may not have taken place during the session. The session injects one final element of beauty into the event (Goodman, 2015) with the reading of a closing poem.

References

Duke Education (n.d.). What is an unconference? https://educationprogram.duke.edu/what-unconference
Educational Transformers UnConference (2017). UnConference schedule. https://educationprogram.duke.edu/what-unconference
Goldberg, D. E. (2009). Pairwork in interdisciplinary and educational initiatives. 2009 39th Frontiers in Education Conference (FIE 2009). IEEE. http://archive.fie-conference.org/fie2009/papers/1561.pdf
Goldberg, D. E. & Somerville, M. (2014). A whole new engineer: The coming revolution in engineering education. Douglas, MI: ThreeJoy Associates, Inc.
Goldberg, D. E. & Somerville, M. (in press). A field manual for a whole new education: Rebooting higher education for human connection and insight in a digital world. Douglas, MI: ThreeJoy Associates, Inc.
Goodman, K. A. (2015). The transformational experience in engineering education. Unpublished PhD dissertation. https://www.proquest.com/openview/7a2969504f1b18f2a3a21a9fc060aee9/1.pdf?pq-origsite=gscholar&cbl=18750

Session 7B: Are ethics guidelines enough? On the effectiveness of ethics guidelines for the design, development, and deployment of conversational agents

Pietro Camin & Samuela Marchiori

As technological advancements permeate every aspect of our daily lives, it has become increasingly difficult to avoid interacting with computer systems. As a result, more and more emphasis is being placed on the importance of creating inclusive systems.

Among such technologies are conversational agents, which enable interactions with technological systems via natural language dialogues.

While not new, contemporary conversational agents differ from their older versions in at least two salient respects. First, they are more ubiquitous than their earlier counterparts. In fact, in addition to being increasingly used to automating mundane tasks in customer support, e-commerce, finance, education, home assistance, and data collection, they have expanded their scope of application to include new domains, such as social work and healthcare (CCNE 2021). Second, limitations that characterized older versions of such technologies are being overcome due to the increased sophistication of new conversational agents, which employ improved Machine Learning techniques (Davenport & Miller 2022).

Despite their numerous benefits, conversational agents are not without their criticisms. Indeed, concerns over their socio-technical implications have been growing together with the popularity of the technology itself (Alnefaie et al. 2021), which was also propelled by competitions funded by big tech companies like Amazon (2017). As a result, several bodies and organizations (most notably: European Union 2016, 2022; HLEG 2019; UNESCO 2021) highlighted the necessity of addressing the ethical concerns raised by conversational agents within the broader AI ethics landscape and advanced guidelines to that effect. Specifically, due to its narrow focus on conversational agents, and the diversity of backgrounds of its contributors, the French National Pilot Ethics Committee's opinion on the ethical issues raised by conversational agents (CCNE 2021) is especially relevant.

In this paper, we focus on ethics guidelines and investigate the extent to which they can effectively safeguard and promote the ethical design, development, and deployment of conversational agents. To this end, we draw on the UNESCO (2021), EU (2016, 2022), HLEG (2019), and CCNE (2022) AI ethics guidelines, which we apply to case studies on systems implementing conversational agents.

References

Alnefaie, A., Singh, S., Kocaballi, A. B., & Prasad, M. (2021). An Overview of Conversational Agent: Applications, Challenges and Future Directions. Proceedings of the 17th International Conference on Web Information Systems and Technologies, 388-396.
Amazon. (2017). Alexa Prize. https://www.amazon.science/alexa-prize (Last accessed 13-01-2023) Davenport, T. H., & Miller, S. M. (2022). Working with AI: Real Stories of Human-Machine Collaboration. MIT Press.
European Commission. (2022). Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts.
European Union. (2016). Regulation (EU) 2016/679.
CCNE. (2021). Opinion: Ethical issues of conversational agents. Adopted on September 15, 2021.
Independent High-Level Expert Group on Artificial Intelligence (HLEG). (2019). Ethics guidelines for trustworthy AI. Brussels: European Commission.
UNESCO. (2021). Recommendation on the Ethics of Artificial Intelligence. Available at: https://en.unesco.org/artificial-intelligence/ethics#recommendation (Last accessed 11-01-2023)

Session 7B: The Ontophobic Turn (Between Philosophy of Technology and Engineering)

Agostino Cera

Inspired by the decades-long work of scholars such Carl Mitcham, this paper deals with a dialogue between philosophy of technology and engineering by sketching a critical historicization of the newest philosophy of technology, starting from the so-called Empirical Turn, according to Hans Achterhuis' definition. My basic assumption is that the philosophy of technology is living an epistemic crisis, in the sense that many prominent scholars in this field show they no longer believe in both "philosophy" (as a peculiar form of knowledge) and "technology" (as something in itself, i.e. an epochal phenomenon). To avoid being deterministic, the current mainstream in the philosophy of technology – especially the postphenomenological approach – has become apologetic, that is no more able to be critical. It is flattened on the status quo, incapable to distance itself from the reality.

The hermeneutical hypothesis at the basis of my historicization is that after 40 years the Empirical Turn proved to be an Ontophobic Turn. By this expression I mean an over-reaction against the transcendental approach to the question concerning technology, in particular against Heidegger's (and continental, more generally) legacy in the philosophy of technology. This over-reaction consists of the transition from an over-distance to an over-proximity: from an disinterest in the empirical/ontic dimension of technology – i.e., its social, political, practical implications – and therefore an over-distance, typical of the first generation of philosophers of technology (i.e. authors as Martin Heidegger, Guenther Anders, Jacques Ellul...), to an almost absolute interest in this empirical/ontic dimension (and therefore an over-proximity), typical of the second generation of philosophers of technology (i.e. authors from Albert Borgmann and Andrew Feenberg to Don Ihde and Peter-Paul Verbeek) with a consequent a priori disinterest in any ontological/transcendental implication of technology. The natural outcome of this attitude is an a priori disinterest in any ontological implication of technology, which is characterised ipso facto as 'essentialist' or 'deterministic' and thus ends up becoming a taboo. That is to say, a real onto-phobia.

Such a rejection culminates in the Mr Wolfe Syndrome, the metamorphosis of the philosophy of technology (particularly the so-called engineering-oriented philosophy of technology, according to Philip Brey) into a positive science which, in turn, depends on an engineerisation/problematisation of reality, i.e. the eclipse of the difference between 'problem' and 'question'.

With regard of this situation, my objection is the following. If "Technology with capital T" becomes nothing, i.e. if the philosophy of technology becomes a problem-solving activity in the presence of concrete problems emerging from the single technologies, then it must be admitted that this activity can be performed much better by specialists (scientists, engineers, politicians...) than by philosophers. As a consequence, the ontophobic turn in philosophy of technology culminates in the disappearance of the reason for a philosophical approach to the question of technology. Given this assumption, the paradoxical accomplishment of the empirical turn should be the final self-suppression of the philosophy of technology.

Session 7B: A Critique of Human-Centred AI: Foucault, Power, and Philosophical Anti-humanism

Mark Ryan

One of the main challenges in debates on the ethics of artificial intelligence (AI) is that we need to develop more sustainable AI. A way to do this is by ensuring the development and use of AI that places the needs and values of human beings at centre stage; often referred to as 'human-centred AI' (HCAI). This approach proposes that there should be a greater level of responsibility and accountability for those producing and using AI, AI models should be sensitive to the diversity of humankind, and AI should be developed to respond to the great sustainability challenges facing humanity now. It is claimed that adopting the HCAI approach will lead to greater social, economic, and environmental sustainability.

HCAI appears to be underpinned by the philosophical approach of humanism. Humanism, and contra HCAI, places human beings as the starting point and centre of all moral inquiry. On first appearance, this appears to be a pragmatic and praiseworthy approach to take because many of the issues and harms resulting from AI are the result of not taking into account the needs and diversity of human beings. However, adopting a human-centred approach to AI may raise some fundamental issues, similar to many of those which were put against humanism in the works of Michel Foucault.

Foucault rejected the idea that humankind is the measure of all things and the humanist conceptualisation of technology aimed to re-orientate technologies towards the human. Foucault claimed that we need to get away from the idea that technology violates or harms a human 'essence', as something that we need to get back to. In this regard, his position would also counter the very premise of HCAI. This presentation will evaluate Foucault's philosophical anti-humanism as a critique of HCAI. It will demonstrate how his 'technologies of power' can be used to understand and resist AI without resorting to the metaphysical claims implied in humanism and HCAI.

Session 7C: Explicit empathy instruction at the intersection of engineering design, technical communication, and philosophy

Katherine Brichacek, Ordel Brown & Laura Pigozzi

Engineers play a vital role in delivering mitigatory effects to wicked problems. However, conventional engineering education relies overwhelmingly on well-structured problems and design challenges and does not prepare students to adequately address wicked problems. Not only are wicked problems daunting and difficult for engineering students, tackling such problems requires unconventional approaches such as an awareness of positionality and sustained empathy in engineering design. While the engineering design process contains the concept of empathy, it is not always explicitly, consistently, and intentionally emphasized.

Drawing on scholarship from the philosophy of empathy and engineering education, we employed a care and virtue ethics-based approach to explicit and intentional empathy instruction in our first-year engineering design curriculum. The curriculum charges students to address wicked problems such as the challenges and inequities faced by people with disabilities. We modified our first-year engineering design thinking course to include activities that encourage students to think of empathy as a skill, a site of learning, and an attitude via engagement with two kinds of empathy: self-oriented and other-oriented. Self-oriented empathy activities promote empathy as a competence that is learned through listening, naming nonverbal cues, and perspective taking. Other-oriented empathy activities enable students to identify and expand their socio-cultural awareness, and grapple with and disrupt biases. More specifically, to encourage consistent, active learning around empathy at different stages of the design process, we created assignments that ask students to reflect on their positionality when tasked with self-oriented and other-oriented empathy exercises. These metacognitive exercises encourage students to recognize the impact of their design choices beyond their own lived experiences.

In the first phase of our research, we used multi-method approaches to capture students' positionality and measure the impact of positionality awareness on shifts in empathetic tendencies. This work-in-progress paper focuses on the initial results from this exploratory phase. They highlight a limited positionality and showed several consistent themes around students' empathetic practices based on perceptions and attitude towards the role and importance of empathy throughout the design process.

Selected References

A. Coplan, "Will the Real Empathy Please Stand Up? A Case for a Narrow Conceptualization," The Southern Journal of Philosophy, vol. 49, 2011. [Online]. Available: https://doi.org/10.1111/j.2041-6962.2011.00056.x. [Accessed October 17, 2022].
S. Fleischacker, Being Me Being You. Chicago: The University of Chicago Press, 2019.
M. Fuller, E. Kamans, M. van Vuuren, M. Wolfensberger & M. de Jong, "Conceptualizing Empathy Competence: A Professional Communication Perspective," Journal of Business and Technical Communication, vol. 35, no. 3, 2021. [Online]. https://doi.org/10.1177/105065
19211001125. [Accessed August 8, 2022].
S. Secules, C. McCall, J.A. Mejia, C. Beebe, A.S. Masters, M.L. Sánchez-Peña, M. Svyantek, "Positionality practices and dimensions of impact on equity research: A collaborative inquiry and call to the community," Journal of Engineering Education, vol. 110, no. 1, Feb. 2021. [Online]. Available: https://doi.org/10.1002/jee.20377. [Accessed: October 24,2022].
J. Walther, S.E. Miller & N.W. Sochacka, "A Model of Empathy in Engineering as a Core Skill, Practice Orientation, and Professional Way of Being," Journal of Engineering Education, vol. 106, no. 1, 2017. [Online]. Available: https://doi.org/10.1002/jee.20301. [Accessed July 26, 2022].

Session 7C: Critically Examining the Broader Implications of Methodological Design in Cross-Cultural, Multi-sited Case Studies of Engineering Ethics Education

Andrea Gammon, Qin Zhu, Rockwell Clancy, Scott Streiner & Ryan Thorpe

Ethics has long been recognized as crucial to responsible engineering, but the increasingly globalized environments present challenges to effective engineering ethics training. This paper is part of a larger research project that aims to examine the effects of culture and education on ethics training in undergraduate engineering students at universities in the United States, China, and the Netherlands. We are interested in how students' curricular and extra-curricular (e.g., internships, service projects) experiences and training impact their ethical reasoning and moral dispositions, and how this differs cross-culturally. To understand this, we are conducting mixed methods research longitudinally over four years to engineering students at our participating universities to gauge their moral dispositions and ethical reasoning skills and to measure any change in these.

This paper, however, is not about the direct outcomes of this research project. Rather, it critically examines our own practices and methods in doing this research. We begin the paper by briefly introducing the larger research project and motivating the use of comparative, multi-institutional case studies as necessary for contextualizing, complementing, and interpreting quantitative data on ethical reasoning and moral dispositions. Because the conditions related to engineering ethics education differ widely per participating institution for institutional (and also likely cultural) reasons, interpreting and analyzing quantitative survey data will require understanding contextual conditions of education at each institution. Comparative case studies can supply missing contextual information to provide a more complete picture of the engineering ethics educational contexts, strategies, and practices at each of the participating universities.

However, in considering how to design and conduct these case studies, we realized we were operating under certain assumptions such as ethics in engineering as separate (and separable from) the 'real,' or technical engineering curriculum. These assumptions have been widely problematized in engineering ethics education (Cech, 2014; Tormey et al. 2015; Polmear et al. 2019); they are assumptions that we in our teaching and research attempt to dispel. Our paper considers (and invites discussion on) the broader implications of methodological design in conducting cross-cultural multi-sited case studies in engineering ethics education research. It explores models for designing and conducting our case studies so as not to reproduce pernicious ideas about social and ethical issues in engineering as subsidiary 'interventions' in the 'actual,' (i.e., technical) curriculum. More generally we discuss how engineering ethics education research methods can be harnessed to overcome this established division.

Works Cited

Cech, E. A. (2014). Culture of Disengagement in Engineering Education? Science, Technology, & Human Values, 39(1), 42–72.
Polmear, M., Bielefeldt, A., Knight, D., Swan, C., & Canney, N. (2019, June 1). Hidden Curriculum Perspective on the Importance of Ethics and Societal Impacts in Engineering Education. https://doi.org/10.18260/1-2--32887
Tormey, R., Le Duc, I., Isaac, S. R., Hardebolle, C., & Vonèche Cardia, I. (2015). The Formal and Hidden Curricula of Ethics in Engineering Education. 43rd Annual SEFI Conference. https://www.sefi.be/wp-content/uploads/2017/09/56039-R.-TORMEY.pdf

Session 7C: Tacking against the headwinds of the military-industrial complex: Engineer's values and professional formation

Rider Foley

Since World War II, the demand for engineers to support national security interests has grown in many parts of the world, especially the United States (US). However, US President Eisenhower's farewell address in 1961 offered a warning about the power of the emergent military-industrial complex. For sixty years engineering graduates have found gainful employment in private and public sector organisations in the military-industrial complex. There are fewer lucrative positions in non-defence industries, thus the prevailing winds push engineers towards careers in defence-related organisations. Yet, some engineer's personal values do not align with national defence and the expansion of military power. This paper explores the question: what norms, values and ethical positions support an engineer's professional formation and how does that relate to defence industries? The values and ethical positions that support decisions to work with (or avoid) defence-related organisations are explored with evidence from reflective writing and interviews with 150 undergraduate engineers. Statements were coded to identify the rationale offered in support or opposition of careers in defence. Engineers that sought careers in the military-industrial complex offered statements on being the best engineer and moral imperative to support national defence (moral imperative), while others offered future earning potential as head-of-household as rationale (paternalism). Those that choose careers in the military-industrial complex reflected on the assumption that if they did not take the job, someone else would (attribution ethics) and that military-based careers were common among their peers (ingroup dynamics). Many had training in Reserve Officers' Training Corps (ROTC) that supports undergraduate education in US universities and aligns with military branches, such as the Army or Navy, which aligned the moral imperative with rule-following behaviour (deontology) and family tradition (familial norms). While the engineers opposed to careers in military and defence industries reflected critically on the misalignment between their personal beliefs and the outcomes of warfare (consequentialism). They understood their earning potential would be lower yet, they recognized they would still find employment opportunities and be financially secure (pragmatism). Many of these engineers were working on military-related project that were required by the university and saw themselves as outsiders or rebelling against the norms of the institution (outgroup dynamics). Those opposed to military careers did not want to be contribute to the deaths of persons in other places of the world, even if they did not know who or when those deaths would occur (care and responsibility). Gary Downey's work on the cultures and philosophy of engineering shed light on the militarization of engineering, and the theory of technopolitics offered by Langdon Winner and Gabrielle Hecht points to the exertion of state-based power through engineered systems. This research explores the tensions between the military-based employment (macro-ethics) and the decisions by individuals to tack against the headwind (micro-ethics). In closing, the ethics of care and responsibility are explored as ways to engage engineers in critically reflective questions about career decisions.

Session 7D: A needs-based theory of social value for research and innovation

Michael J. Bernstein, Lauren Withycombe Keeler, Luke Boyle & John Harlow

Conventional arguments for public funding of research and innovation focus on addressing market failures to generate economic value. In many cases, market interventions targeting economic value come at the expense of broader public value. Other times, R&I market interventions usurp economic and public value alike in service of generating "science value" -- pursuit of knowledge products independent of economic or public use.

Each of these three frames -- economic value, public value, and science value -- seek to maximize singular facets of benefit. Economic value prioritizes profit maximization; public value a notion of a democratic well-being; science value the progression of knowledge. Most people—our lived experiences, wants, needs, and values—are absent from such frames. Absent also are, frequently, the concerns of ecosystems and more-than-human beings. Paradoxically, scientific and technological innovation in these three frames makes promises about human, environmental, and societal benefit, yet never seems to ask what people might want or need, and correspondingly, what might be of broader "social value."

Diving into conceptualizations of social value across fields reveals clear agreement that the concept has something to do with needs and wellbeing beyond that of the individual. Yet abstractions like societal wellbeing or societal needs are rarely made concrete. Current claims of "social value," too, are disconnected from human needs or their satisfaction. While many fields cite psychological research on proposals of universal human values (e.g., Schwartz 1994), few operate with a theoretically grounded definition of social value (Lashitew et al. 2021).

In this paper, we draw on the human-scale development model (H-SD)(Max-Neef 1992), and public value theory (Bozeman 2007) to offer a need-based theory of social value. We review how "human needs" might be defined, differentiated for individuals and societies, and theorized in relationship to modes of human existence and society. The H-SD distinction between needs and modes of satisfaction resonates and adds to conversations of value change more generally (van de Poel 2022). Grounding a definition in human needs offers a way to empirically investigate social value for research, innovation, business, public administration, and community action. Our intention is not to assert truth claims about what all human beings must need. We intend rather to offer a theoretically sound basis to inform research, innovation, policy, and practice in a manner more faithfully approximating genuine contributions to the kind of broad human wellbeing so often absent in predominant economic, public, and scientific value frames.

References

Bozeman, B. (2007). Public values and public interest: Counterbalancing economic individualism. Georgetown University Press.
Lashitew AA, Narayan S, Rosca E, et al. (2021) Creating Social Value for the 'Base of the Pyramid': An Integrative Review and Research Agenda. Journal of Business Ethics1–22.
Max-Neef, M. (1992). Development and human needs. In P. Ekins & M. Max-Neef (Eds.), Real-life economics: Understanding wealth creation (pp. 197–214). Routledge.
Schwartz, S. H. (1994). Are There Universal Aspects in the Structure and Contents of Human Values? Journal of Social Issues, 50(4), 19–45.
van de Poel, I. (2022). Understanding value change. Prometheus, 38(1). https://doi.org/10.13169/prometheus.38.1.0007

Session 7D: The Gap in Value Change

Michael Klenk

It seems obvious that values change. Two hundred years ago, the US economy thrived on slavery; millions of Chinese girls' feet were bound, children were working long days in factories, and English gentlemen were killing each other in duels; women's right were nowhere to be found; animal rights were unheard of. Today, civil rights are solidifying, and the idea that animals ought to be treated humanely is gaining popularity. Foot binding and dueling are recognized for the bizarre cruelties that they are. In many places, child labor has long been abolished.

Though philosophers have since Plato tried to understand the notion of value, they have only recently begun to theorize about value change (cf van de Poel 2018). While important contributions have been forthcoming in recent years, there is no unified theory yet and, perhaps more relevant, no consensus about how such a theory should look like and what kind of phenomena it ought to explain to be considered a viable candidate theory.

My aim in this talk is to introduce and defend a key desiderata for any philosophically viable theory of value change.

I will argue that any theory of value change must account for what I call "the Gap." The gap is a conceptual and normative distinction between 'values' understood as mind-dependent (biological, sociological, or psychological) supervenient properties (the subjective sense) and 'values' as prima facie mind-independent supervenience bases of such properties (the objective sense).

First, I show that there are conceptual differences between values in the subjective and values in the objective sense. The relation between subjective and objective values is not one of implication or causation but one of grounding, and I will explain that relation in some detail.

Second, I argue that there is a crucial normative difference between subjective and objective values. The latter, but not the former, give us reasons to act, desire, and feel in certain ways.

The aim for my account of 'The Gap' is to account for pertinent metaethical distinctions without overburdening the debate about value change with its practical aspirations e.g. for value sensitive design. As such, my account of the Gap is explicitly understood as a model and my aim in the presentation is to evaluate it in terms of accuracy and fruitfulness.

Session 7D: Adding to understanding value change with Luhmann's event-concept

Gunter Bombaerts

Value change in sociotechnical systems recently received increasing attention [1]. This led to a value change taxonomy [1] or a conceptual framework based on John Dewey of value change manifestations [2]. This contribution wants to add to the further understanding what value change means in sociotechnical systems.

I will use Niklas Luhmann's theory here and am aware I describe it very succinctly.
According to this theory, individuals are faced with the world's complexity, which creates contingency ("also being possible otherwise")[4:25]. Expectations narrow down the possibilities [4:96] and expectations of expectations create structures that become norms in social systems [4:292]. Based on John Dewey and Floyd Allport, Luhmann defines "event" as the "smallest possible temporal atom, an indivisible, all-or-nothing happening" [4:287]. In an event, a selection is forced [4:60] and the direction that is forced creates an indication. "What is possible is interpreted as the difference between different potentialities (including the one that is presently actualized and to which one can return), and the possibility of being actualized is then indicated in its identity as "this-and-not-something-else."" [4:66] In such an event, "every complex state of affairs is based on a selection of relations among its elements, which it uses to constitute and maintain itself. The selection positions and qualifies the elements, although other relations would have been possible." [4:25] Values are "general, individually symbolized perspectives which allow one to prefer certain states or events." [4:317] and fall under Luhmann's notion of relations that influence the positioning.

This very selective set of elements of Luhmann's theory has several implications for value change.
• Values determine the positioning in an event and the positioning the values in a next event. It could be a key to understand the interaction between value change and technology innovation.
• Luhmann's states that "values are in force (gelten) in the communication mode of the assumption. It is presupposed that there is consensus on value appraisal"[3]. Values and other event elements are used in events to make a difference, but also to socialize [4:241] and to determine stability (structure [4:307]).
• From a Luhmannian perspective, value change is more fundamental than value stability. Value change is "the normal" as a reaction to a contingent and complex world and meaning and information are based on "a difference that makes a difference" [4:242]). Value stability is the exception, as it can only emerge as a (temporally) structure.

These are a few consequences of a (too?) concise explanation of Luhmann's theory. I am looking forward explaining more at the conference.

References
[1] I. van de Poel, "Design for value change," Ethics Inf Technol, vol. 23, no. 1, pp. 27–31, Mar. 2021, doi: 10.1007/s10676-018-9461-9.
[2] I. van de Poel and O. Kudina, "Understanding Technology-Induced Value Change: a Pragmatist Proposal," Philos. Technol., vol. 35, no. 2, p. 40, Jun. 2022, doi: 10.1007/s13347-022-00520-8.
[3] N. Luhmann, Gibt es in unserer Gesellschaft noch unverzichtbare Normen? Müller, Jurist. Verlag, 1993.
[4] N. Luhmann, Social systems. Stanford University Press, 1995.

Session 8 (Keynote): On the nature of values and their role in IT system design

Sarah Spiekermann

This talk gives an insight into the nature of values, how the might be properly defined in the light of phenomenological research and Material Value Ethics. It then illustrates why an appropriate understanding of values is necessary to systematically derive IT system requirements to build „ethical" or simply „good" systems. The talk then illustrates how this knowledge is used in „Value-based Engineering" with ISO/IEC/IEEE 24748-7000 (short IEEE 7000).

Session 9A: Adversarial Machine Learning for Deliberating Decision-Making in Elections

Syafira Fitri Auliya

The advance in Artificial Intelligence (AI) is a double-edged sword, with its biggest social cost is possibly the erosion of our democracy and privacy. Recent years have suggested practical examples of how democracy and privacy are directly interrelated when AI is concerned, namely, how AI affects the deliberate decision-making of the people. In that context, AI has a history of influencing people's choices in elections using unauthorized use of data, reducing people's trust in democratic institutions, and becoming a big hit to the core values of democracy itself. Because of these previous negative experiences, it is understandable that, predominantly, related research initiatives are focused on solving the negative influences of AI on the issue. Nevertheless, the philosophy of technology provides more holistic and complex ideas on how to engage technologies. Specifically, depending on how we conceptualize AI philosophically, AI can also be conceived of as a mediator in deliberate decision-making. Looking at it as such would allow us to have a more positive outlook on the potential of AI in making better deliberate decision making, with its close interrelations to privacy and democracy.

While a sole technology is impracticable to solve the complex issues of democracy and privacy, exploring the possibility of technology can help us be more reflective regarding its effects. In this presentation, I propose one way to cautiously explore AI's potential and how it can enable better deliberate decision-making of people. More specifically, I focus on the privacy of deliberation in the election, especially with individual data from social media. The issues of privacy affecting elections will continue to be relevant in the coming years because countries will keep having elections. Negative uses of AI to manipulate the election by keeping the voters' privacy at threat will risk the stakes of fair election processes.

One of the nascent areas in this direction explored in this presentation is the development of Adversarial Machine Learning (AML). AML is a rising issue in the AI community, known as the possibility of fooling the AI system by taking advantage of the characteristics of deep networks used in AI systems to help large-scale process data and make seamless decisions (Szegedy et al. 2014). Here, I introduce the novel possibility of AML playing a role in protecting people's privacy in the election process by fooling other malicious AI system that tries to profile and influence users into specific agendas. Helping to facilitate data privacy on social media, I expect AML can help to facilitate the deliberate decision-making of people in elections.

References

Szegedy, Christian et al. 2014. “Intriguing Properties of Neural Networks.” http://arxiv.org/abs/1312.6199 (October 3, 2022).

Session 9A: Autonomy in the test of low-tech devices

Aurélien Béranger & Hugues Choplin

Over the last ten years in France, we have seen the rise of a movement, bringing technological alternatives back into fashion from an ecological perspective. This low-tech movement, heir to a long technocritical tradition (Jarrige 2014; Winner 2022), is based on the articulation of ecological considerations and a project of technical democracy (Callon, Lascoumes, and Barthe 2014) through a discourse on autonomy, in the sense given by Ivan Illich (2014). However, this articulation does not seem to us to be self-evident, as the discourses on autonomy reconduct a form of anthropocentrism a priori contrary to the ecological thought (Bourg and Fragnière 2014). In this paper, we propose to problematize the link between autonomy and low-tech, with an approach that crosses philosophy and Science and Technology Studies, informed by participant-observation field surveys conducted on the low-tech movement. The devices we will focus on are those that participate in the "dream of disconnection" (Lopez 2014) by seeking to break away from dependencies on large technical systems. We will confront them with relational conceptions of autonomy, wether it relates to the idea of relational autonomy (Mackenzie and Stoljar 2000), to care ethics (Tronto, 2009) or to feminist theories of subsistence (Pruvost 2021). These perspectives on autonomy generally suggest a recognition of human interdependence. What happens when this interdependence is no longer only to others but to more-than-human elements, as is the case with some low-tech devices that mediate air or atmosphere, water or solar radiation? Is this a new form of relational autonomy? Or, as Levinas (2009) suggests, a singular relationship to the elements that thwarts the autonomy/ heteronomy dualism?

References

Bourg, Dominique, et Augustin Fragnière. 2014. La pensée écologique: une anthologie. L’écologie en questions. Paris: Presses universitaires de France.
Callon, Michel, Pierre Lascoumes, et Yannick Barthe. 2014. Agir dans un monde incertain: essai sur la démocratie technique. Éd. révisée. Points 734. Paris: Points.
Illich, Ivan. 2014. La convivialité. Paris: Éditions Points.
Jarrige, François. 2014. Technocritiques : du refus des machines à la contestation des technosciences. Hors collection Sciences Humaines. Paris: La Découverte.
Lévinas, Emmanuel. 2009. Totalité et infini: essai sur léxtériorité. Livre de Poche Biblio essais 4120. Dordrecht: Kluwer Academic.
Lopez, Fanny. 2014. Le rêve d’une déconnexion: de la maison autonome à la cité auto-énergétique. SC. Paris: Éditions de la Villette.
Mackenzie, Catriona, et Natalie Stoljar. 2000. Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self. Oxford University Press.
Pruvost, Geneviève. 2021. Quotidien politique: féminisme, écologie et subsistance. L’horizon des possibles. Paris: La Découverte.
Tronto, Joan C. 2009. Un monde vulnérable: pour une politique du care. Traduit par Hervé Maury. Textes à l’appui. Paris: Éditions la Découverte.
Winner, Langdon. 2022. La Baleine et le réacteur: A la recherche de limites au temps de la haute technologie. Traduit par Nicolas Casaux. 2e éd. [1986]. Technocritique. Herblay: Editions LIBRE.

Session 9A: Ethical, Legal, and Socio-Economic Analysis of Agrifood Data Marketplaces - Lessons from FlexiGroBots

Artur Bogucki & Paula Gürtler

The digital transformation of the agricultural sector brings to the fore questions on the value of agri-food data and the key role data spaces and agricultural data marketplaces play in scaling up the adoption of the new technology [1]. These questions are met by concerns over creating dependencies of farmers on agricultural technology providers (ATP) and by proposals on how such pitfalls can be avoided. A growing body of literature aims to address this risk by developing technological requirements and recommending design principle for data spaces. For example, the Position Paper of the International Data Spaces Association (IDSA) recommends four design principles: Data Sovereignty, data level playing field, public-private governance, and decentralised soft infrastructure [2].

In our presentation, we make a case for combining these technical design principles with an ethical theory. The added value of the ethical theory is that it enables coherent interpretation of the technical principles in their implementation into data governance structure. We argue that the four design principles recommended by IDSA can be strategically supported by a theory of relational egalitarianism: “Egalitarians aim to replace social hierarchies with relations of social equality” [3]. A theory of relational egalitarianism in conjunction with theories on social welfare provide a basis for coherent criticism and a potential roadmap for designing a remedy for the issues of market centralization, information asymmetry, and a sector wide anti-commons effect [4].

Illustrating the utility of introducing an ethical theory to the debate on data spaces, we, the Centre for European Policy Studies (CEPS), are applying these ethical findings to our work on the Horizon 2020 FlexiGroBots project. At the fPET 2023, we would like to present part of our research on the optimal structure of agricultural data governance and data spaces. This optimal structure brings together insights from moral philosophy, law, and behavioural economics. With our research, we develop a framework which enables the application of socio-economic tools to agricultural data spaces and market places, by assigning economic value to data. Our work also involves the formulation of a set of contracting standards between the owners of agricultural data and agricultural technology providers (ATP). The objective of the contractual standards is to provide legal certainty to contracting parties, while also facilitating the values of equality and social welfare.

CEPS’ recommendations for data space governance structures are aimed at helping to establish a sustainable and efficient agri-food data marketplace, with lessons transferable beyond the agricultural sector.

References

[1] B. Otto, M. ten Hompel, and S. Wrobel, Designing Data Spaces. Cham: Springer International Publishing, 2022. doi: 10.1007/978-3-030-93975-5.
[2] L. Nagel and D. Lycklama, ‘Design Principles for Data Spaces’, Apr. 2021, doi: 10.5281/ZENODO.5244997.
[3] E. Anderson, Equality. Oxford University Press, 2012. doi: 10.1093/oxfordhb/9780195376692.013.0002.
[4] D. Hunter, ‘Cyberspace as Place and the Tragedy of the Digital Anticommons’, Calif Law Rev, vol. 91, no. 2, p. 439, Mar. 2003, doi: 10.2307/3481336.

Session 9B: Beyond reliability: values and Trustworthy AI

Daniele Chiffi, Mattia Petrolo, Viola Schiaffonati & Giacomo Zanotti

The notion of Trustworthy AI (TAI) has recently gained much attention. Moreover, it plays a central role in the European ethics-based effort to regulate AI (AI HLEG, 2019). However, some authors have argued that the concept of trust is problematic in its application to AI systems (Hatherley 2020; Ryan, 2020). The debate largely boils down to the choice between two alternatives. Opting for a cognitive account of trust, that takes trust to be a matter of purely rational choice and probability estimation, allows us to make sense of the notion of TAI but makes it hard to differentiate TAI from merely reliable AI — a distinction usually deemed essential. Motivational accounts of trust, instead, focus on the trustee’s motivations and moral obligations and provide a clear distinction between trustworthiness and reliability. However, since AI systems lack motivations and moral obligations, the notion of TAI turns out to be a categorical error.

In both cases, the notion of TAI somehow reduces to the one of reliable AI. We contend that this outcome is undesirable. As widely acknowledged in the literature, AI is highly value-laden (see, e.g., Biddle, 2020). The notion of TAI plays a crucial role in this respect, allowing us to go beyond mere reliability and capture the ethical and value-related dimensions involved in the design and use of AI systems.
We provide an alternative framework for addressing the question of TAI. In our view, the current debate builds upon two wrong assumptions, namely that (i) trust in AI should be uncompromisingly modelled on interpersonal trust and (ii) the attribution of trustworthiness to AI systems should be understood literally. We argue in favour of a more pragmatic stance that takes talk of trustworthy AI to be an instance of loose talk, in which the communicated content of the utterance differs from its literal content (on the dynamics of loose talk, see Carter, 2021). In particular, we maintain the focus on the ethical and value-related dimensions of trust without literally ascribing interpersonal trustworthiness to AI systems. This way of framing the discourse on TAI allows us to go beyond merely reliable AI without making categorical mistakes.
We conclude by extending our pragmatic approach to the question of the determinants of trust in AI. The philosophical literature typically conceives trust in an idealized way, as resulting from the obtaining of some fixed and clearly identified factors, and the debate on TAI largely inherits this tendency. We argue that a different approach should be adopted, focusing on the role played by contextual factors.

References

AI HLEG (2019). Ethics Guidelines for Trustworthy AI.
Biddle, J. B. (2022). On predicting recidivism: epistemic risk, tradeoffs, and values in machine learning. Canadian Journal of Philosophy, 52(3), 321-341.
Carter, S. (2021). The dynamics of loose talk. Noûs, 55(1), 171-198.
Hatherley, J. J. (2020). Limits of trust in medical AI. Journal of medical ethics, 46(7), 478-481.
Ryan, M. (2020). In AI we trust: ethics, artificial intelligence, and reliability. Science and Engineering Ethics, 26(5), 2749-2767.

Session 9B: Data Quality as an Emerging Issue for AI: The Case of Wearable Technology

Stefano Canali, Viola Schiaffonati & Andrea Aliverti

The contemporary debate on the ethics and epistemology of Artificial Intelligence (AI) has focused on various issues affecting the data used for representing populations, training algorithms, and testing models, such as bias and reproducibility. What has not been discussed as extensively is data quality and its impact on the performance, trustworthiness, and explainability of AI models. Data quality is a key epistemic concern in the sciences, where several data practices are aimed at checking and assessing the quality of data produced and used in research and has been discussed in the philosophy of data-intensive science and data studies. In this paper, we present the need for more focus on data quality by analysing current applications of AI models in wearable technology and we provide a framework for conceptualising data quality as an emerging issue in the philosophy of AI.
Wearables are devices that are worn on our bodies and collect large volumes of data on individual health. As such, wearables are increasingly used by the general population and biomedical researchers and practitioners, thanks to their abilities for constant and remote monitoring and individual recommendations and interventions. Wearables are also considered a key source of health data to train algorithms, which are currently used to e.g. estimate physical activity, evaluate sleep quality, and detect COVID-19.
Despite the significant promises of wearables and AI models deployed in this context, we argue data quality is an emerging concern in connection to three main issues. First, the variability of devices, sensors, approaches to data collection and analysis make it difficult to create baseline standards of data quality. Second, the ways in which data are collected, analysed, interpreted by algorithms are usually not accessible and transparent, a crucial issue when assessing data quality. Third, we raise questions on representativity, as data collection tends to over-represent younger and wealthier users and exclude significant groups and minorities, thus leading to datasets that often fail to represent the general population.
On the basis of this case study, we present a framework to conceptualise and approach data quality as an emerging issue for the philosophy of AI and ML. First, we argue that data quality should be considered an epistemic issue with substantial ethical implications. For instance, policy-making based on non-representative data leads to public health that is not ethically distributed and socially acceptable. Second, we present approaches that should be grounded on both epistemic and ethical considerations. For instance, focusing exclusively on the accuracy of measurement does not necessarily help to solve issues of representativity, as barriers of access and inclusion remain.
We thus frame our work as a contribution to ongoing discussions on the ethics and epistemology of AI, at the intersection of philosophy and technology, as a result of interdisciplinary collaborations between philosophers of science and technology and biomedical and computer engineers.

Session 9B: Manipulating the scaffolded user

David Spurrett

One of the key concerns in the ethics of artificial intelligence is the deployment of AI for the manipulation of users or their behaviour. What we count as manipulation, and what means we regard as possible to achieve it, depends partly on our view of how agency and cognition work more generally. Here I focus on manipulation of behaviour taking as a starting point the view that human cognition extensively scaffolded. Cognition is scaffolded when it relies in significant ways on structure and resources external to the skull and skin. Human cognition is distinguished in part by the magnitude of our reliance on scaffolding, and by the variety and sophistication of the forms of scaffolding we have developed and transmit culturally. Most arguments supporting the idea that cognition is scaffolded focus on benign cases, which augment or support the agent attempting some task (e.g. Clark 2003). But scaffolding needn't be benign. Here I make two arguments:

First, that standard conceptions of how users can be manipulated in both technology ethics and policy tend to presuppose either a 'Humean' or 'Behavioural Economic' agent. The Humean agent is an actor defined by a set of beliefs and desires, and who can be manipulated by interfering with either of those, typically by manipulating their beliefs through misrepresenting the world, or associations between world states and expected desire satisfaction (think of advertising). The behavioural economic agent, on the other hand, is characterised by the familiar heuristics and biases, and possibly manipulated by their differential triggering or nudging (Thaler & Sunstein 2008). Although both are fruitful approaches, neither can adequately accommodate the scaffolded agent, and that agent's distinctive vulnerabilities. The scaffolded agent, precisely because they rely on external structure for their cognitive performance, can be manipulated by changing things in the world.

Second, that there are already real examples of manipulation and exploitation of scaffolded agents. These are cases where the user's cognitive processes depend in significant ways on external structures, but these structures don't reliably serve the interests of users. I'll use features of electronic gambling machines and casino management systems as concrete examples of this 'Hostile Scaffolding' (Spurrett & Timms, MS). While these might seem like cases from a narrow and unusual niche, many of the same design features are at work on social media and other virtual environments, where they are deployed to extend time on device or manipulate engagement, rather than to encourage extended gambling.

Having established those points, I extract a few practical recommendations for the conduct of engineering and technology, including showing how some existing policies fail to articulate the vulnerabilities of the scaffolded user, and so leave legal room for forms of exploitation that depend on scaffolding.

References

Clark, A. (2003) Natural Born Cyborgs. OUP.
Thaler, R.H., & Sunstein, C. (2008) Nudge: Improving decisions about health, wealth and happiness. Penguin.
Timms, R. & Spurrett, D. (Under review) Hostile Scaffolding.

Session 9C: A virtue ethics account of moral responsibility

Zoë Robaey

In this paper, I explore the conceptual short-comings of the notion of responsibility-as-virtue when it comes to understanding responsibility in practice. Conceptually, Williams (2008) speaks of responsibility as virtue as "the readiness to respond to a plurality of normative demands", and Sand and van de Poel (2018) in the context of responsibility in innovation as the "willingness to take on responsibility". This paper investigates two main short-comings of these accounts for responsibility in practice. First, they are non-prescriptive of what responsibility should entail as it leaves much open to the responsible agent's decision or interpretation. Second, these accounts speak of the role of institutions for responsibility without specifying what this role is besides supporting responsible agents. These short-comings are quite critical as they provide little guidance to either individuals or institutions in how to realize more responsible practices.

Building on these interesting yet insufficient accounts on the link between responsibility and virtue ethics, I ask: How can virtue ethics enrich a concept of responsibility in practice?
In order to answer this questions, I develop three suggestions. A first suggestion I make is to consider groups of people instead of institutions, inspired by Astola (2022). However, instead of investigating blame or praise, I look at forward-looking moral responsibility and suggest that we should consider responsibility as an outcome of other desirable actions. This leads to my second suggestion: since responsibility is an outcome, we need a certain set of virtues that help sustain good relationships between people and with things in the context of morally relevant issues. Finally a third suggestion is that these relationships cannot be sustained by an agent alone, and I therefore expand on Sie's account of shared moral responsibility (2018).

Last but not least, in order to illustrate how this account is in fact practical, I reconstruct my experience as an embedded ethicist within an innovative biotechnology start-up, which informed the formulation of this virtue ethics account of moral responsibility.

References

Astola, M. (2022). Collective Responsibility Should be Treated as a Virtue. Royal Institute of Philosophy Supplements, 92, 27–44. https://doi.org/10.1017/S1358246122000133
Sie, M. (2018). Sharing responsibility: The importance of tokens of appraisals to our moral practices. In M. Oshana, K. Hutchison, & C. Mackenzie (Eds.), Social dimensions of moral responsibility (pp. 300–323). Oxford University Press.
van de Poel, I., & Sand, M. (2021). Varieties of responsibility: Two problems of responsible innovation. Synthese, 198(19), 4769–4787. https://doi.org/10.1007/s11229-018-01951-7
Williams, G. (2008). Responsibility as a Virtue. Ethical Theory and Moral Practice, 11(4), 455–470. https://doi.org/10.1007/s10677-008-9109-7

Session 9C: Defining Ethical Risk for Technology

David M. Douglas & Justine Lacey

'Ethical risk' is increasingly used in discussions of the ethics of artificial intelligence (AI) and other new and emerging technologies. However, the term itself is rarely defined, obscuring what distinguishes 'ethical risk' from other forms of risk, such as reputational or legal. Existing accounts of ethical risk, such as those of the British Standards Institute (BSI) or by Bertrand Andre Rossert, either conflate ethical risk with the likelihood of compromising psychological, social, or environmental wellbeing (BSI), or describe it as the possibility that some individuals will be unnecessarily excluded or instrumentalised without committing to a specific theory of justice to determine when such exclusion or instrumentalisation is unjust (Rossert).

We present an account of ethical risk for technology that centres on the interconnection of risk and responsibility. We define ethical risks as technical risks connected with a technology that may cause stakeholders to fail one or more of their ethical responsibilities. Technical risks are characteristics of a technology that may result in unwanted events. The stakeholders connected with a technology have both external ethical responsibilities that follow from being moral agents, and internal ethical responsibilities that follow from their social or occupational role as a stakeholder. These internal responsibilities are forward-looking ethical responsibilities (such as obligations, accountability, and blameworthiness) as they are concerned with preventing or mitigating potential ethical risks that may occur.

Mapping the process within which a technology is designed and used identifies the stakeholders connected with the technology, their roles within the process, and the implicit and explicit connections between stakeholders themselves. This process begins with the gathering resources necessary to develop the technology, the decision to use the technology, the use of the technology itself, and the evaluation of its effectiveness. The value of mapping this process is that the roles and connections between stakeholders identify the internal ethical responsibilities specific to the technology. Subsequently, the connections between stakeholders can be evaluated using Sven Ove Hansson's ethical risk analysis (eRA) approach to determine how the benefits, costs, and the ability to make decisions about a technical risk are distributed between them. Where the ability to make decisions about a technical risk is separated from gaining the benefits of or being exposed to the costs of that risk (or both), the affected stakeholders are dependent on the decision-maker.

Building on this work and drawing on Frank Lovett's theory of domination and justice, such dependencies may create a vulnerability to the arbitrary power of the decision-maker to influence the technical risks that affect other stakeholders. The ethical responsibilities of the decision-maker are the constraints that should prevent the decision-maker's arbitrary power over the technical risks of a technology from being used to dominate others, and the key ethical risk for the decision-maker is the failure to recognise and follow these constraints. This account of ethical risk describes how it differs from other forms of risk, and connecting ethical risk with responsibilities that constrain arbitrary power over technical risks provides a clearer link between ethical risk and justice.

Session 9C: Defining Ethical Risk for Technology

David M. Douglas & Justine Lacey

In recent years, professional societies, scholars and educators, and corporations within the fields of engineering and computer science have been grasping for better principles and frameworks for the ethical and social responsibility that engineers, computer scientists, and technologists should embody. Many professional societies (e.g. NSPE, IEEE, ACM) have revised their codes of ethics. Scholars and educators have launched numerous research projects and educational experiments to determine the values and ethical competencies professionals should learn and apply. Corporations have publicized their ethical commitments and formed consortia to govern ethical approaches to research and development areas like artificial intelligence.

We know there are fundamental problems. In 2014, Erin Cech identified a "culture of disengagement" in engineering that weakens engineering students' commitments to public welfare during their undergraduate years. She argued that ideologies of 1) depoliticization, 2) technical/social dualism, and 3) meritocracy were key pillars of engineering's disengaged culture. These ideologies undermine engineers' social responsibility by positioning technical expertise as supremely relevant and perceiving existing social, economic, and political structures as fair and just. In subsequent studies, Cech and co-authors call for engineering education and epistemologies of engineering that repoliticize the profession and its work. I believe civic professionalism answers that call.

Based on Harry Boyte's concept of "citizen professionalism" and Albert Dzur's concept of "democratic professionalism," civic professionalism is both a professional identity, anchored by civic attitudes and related values, and a set of normative professional practices that rely on civic knowledge, skills, and habits, which augment specialized technical competencies. In contrast to typical outside experts, Boyte argues citizen professionals see their role as co-creators and facilitators of problem-solving. They acknowledge that they too are citizens alongside many other stakeholders and should share power over decisions. They acknowledge the limits of their knowledge and expertise in different contexts and embrace local knowledge. Their work serves the common good by solving technical problems while also building and strengthening relationships.

Similar to Boyte's description, Dzur's democratic professionals resist the technocratic urge to flatten complex problems into challenges well-suited to professional methods beyond the lay public's reach, and instead create space for deliberation and collective action regarding social and political issues beyond the borders of their own professional domains. Dzur specifically offers democratic professionalism as a middle ground between a "social trustee" model of professional and its radical critique, which would seek to deprofessionalize expertise and recover all such power for the public. Dzur argues that professionals and laypeople both have a stake in professional decisions and should share oversight of professional ethics democratically.

This is where engineering and computer science should head—framing social and ethical responsibility in terms of a broader civic and democratic responsibility. Although civic professionalism does not have a monopoly on these tenets, which we can locate in other popular ethical engineering frameworks, it does provide a compelling emphasis on epistemic humility, politics, and the common good with clear pedagogical opportunities as this paper will show.

Session 9D: Biorobots as objects, tools or companions? An ethical approach to understand bio-hybrid systems

Rafael Mestre & Anibal Monasterio

Great challenges are expected with the progressive incorporation of artificial systems and in particular robots in society. Robotic technologies have advanced quickly in the last few decades. A paradigm shift in robotics may result from the integration of biological tissue or cells with artificial parts, or “biohybrid robotics.” Many ethical problems have arisen with these breakthroughs and new paradigms. A set of ethical concerns pose the following question: are these biorobots best viewed as objects, tools or companions? Using the Actor Network Theory (ANT), an approach that sees networks of relations between objects, discourse and humans (where non-humans also participate in the social world), and incorporating a normative sense to this approach, we enrich the ANT theory with an ethical point of view in order to understand that biorobots by their very nature call into question the dichotomy between/natural and artificial/object and tool/companion and human. This communication examines some of the science behind bio-hybrid systems before going beyond what is now possible to investigate philosophical and ethical issues on how bio-hybrid systems may be viewed to affect the determination of human relationships. We end this communication with some ideas on how to proceed with this understanding of bio-hybrid systems while taking into consideration the moral and philosophical concerns.

Session 9D: Machine Mediated Gestation - Ontological Distinctiveness and Normative Implications

Llona Kavege, Julia Hermann & Olya Kudina

Research developments in biomedicine suggest that artificial wombs could become a reality. More accurately described as ectogestative technologies (EGTs), they will enable the partial gestation of human fetuses in an ex utero device beyond the maternal body. Researchers aim to improve the viability of extremely premature newborns in neonatal intensive care by mirroring conditions in utero to treat them as if they had never been born. The promise of this technology has raised considerable debate in the media and academic literature. One particular point of contention is how to adequately refer to the gestating human in the device. If a premature fetus is ‘born’, meaning delocalized from its mother’s womb and translocated into an ex utero device, then is it still a fetus, a neonate, or something altogether different? In this paper, we argue for the ontological distinctiveness of the human gestating in an EGT and propose a novel conceptualisation of it, exploring legal, social, and moral implications.

EGTs will consist of a sterile hermetically sealed pouch with a close circuit artificial amniotic fluid, and a cannulation system that will connect the umbilical cord to an outside device and serve as an artificial placenta. The circuit will be pumpless and solely powered by the fetal heart to replicate in utero blood circulation. In the existing ethical and philosophical literature, several terms for the human being gestating ex utero are circulating. E.C. Romanis, one of the first to advocate the ontological distinctiveness of ectogestating entities, introduced the term ‘gestateling’. Research teams in the United States and in the Netherlands have also introduced new terms, such as ‘fetonate’ and ‘perinate’ respectively. The lack of a consensus and disparate debate demonstrates how EGTs are disrupting basic concepts related to the beginning of life. Because of the biomedical context and sensitive nature of this technology, the emerging limitations of existing concepts will have direct practical implications in research development and design, as well as in the legal, moral, and social spheres. On the surface this might seem like semantic trifle, however, the ontological status of the human constituted in its interaction with the EGT matters because it will shape our moral intuitions regarding its moral status, and how it will be interpreted and interacted with by others.

Consequently, this paper aims to use EGT as a case study to demonstrate the limitations of existing concepts when faced with emerging technologies. We will critically engage with the existing literature and terms for the ectogestating human being and demonstrate why current concepts are limited. Thereby, we will draw from the notion of technological mediation and recent scholarship on the metaphysics of pregnancy to reframe the discussion. Next, we will enrich the debate by introducing a novel perspective to categorize the subject of ectogestation: the gestateling as a cyborg. Finally, we will conclude by reviewing present challenges to the development of EGTs, addressing the politics of conceptual engineering, and driving down the point that characterizing the subject of ectogestation is at its core a normative issue.

Session 9E: Training engineers for sustainability, but which one? A discussion on alternatives to the “good Anthropocene”

Hugo Paris, Nicolas Freud & Caroline Ladage

Our submission explores how sustainability is and could be introduced into engineering curricula. Based on a study case of a curricular reform at the Lyon National Institute of Applied Sciences (INSA Lyon, France), we discuss the modern heritage of French engineering schools and how it can frame the understanding of sustainability toward a “good Anthropocene”. We then offer some reflections on this perspective and try to propose a critical approach for educating engineering students for sustainability.

Given the strong links between French engineering schools and positive philosophy (1) and more generally with modernity (2), this educational field carries an important epistemological heritage based on a realist ontology and a nature-culture divide. The development of a holistic science of sustainability supported by the UNESCO (3) is one of many examples suggesting that the modern paradigm alone may not be adapted to solve the “wicked problems” (4) of the Anthropocene. However, there is no consensus on this position and some researchers have argued for a “good Anthropocene” perspective based on a new wave of modernization in which humans would extend their control over planetary regulation (5). In this communication, we stress that this perspective can lead to a renewal of what Foucault called pastoral power, as shown by Paul Crutzen (6) statements implying that engineers and scientists should “guide mankind towards global, sustainable, environmental management”. Other narratives are possible, e.g. as proposed by Bonneuil (7).

From this point of view, we discuss how a pluralist approach to the role of engineers in the Anthropocene could be elaborated supporting new forms of democratic structures and organizations, as Latour’s proposal of a “democracy extended to things themselves”(2). Such democracy implies to reconsider the relations between humans, other living beings and things in order to build new political communities able to transform the world in a sustainable way.

Reference

(1) Jamison, A., Kolmos, A., & Holgaard, J. E. (2014). Hybrid Learning : An Integrative Approach to Engineering Education. Journal of Engineering Education, 103(2), 253 273. https://doi.org/10.1002/jee.20041
(2) Latour, B. (1991). Nous n’avons jamais été modernes, La Découverte.
(3) UNESCO (2017). Guidelines on Sustainability Science in Research and Education. https://unesdoc.unesco.org/ark:/48223/pf0000260600
(4) Keenan, W. J. F. (2020). Learning to survive : Wicked problem education for the Anthropocene age. Journal of Global Education and Research, 4(1), 62 79. https://doi.org/10.5038/2577-509X.4.1.1038
(5) The notion of « good Anthropocene » was coined by Erle Ellis and set out in a 2015 collective text supported by The Breakthrough Institute : Asafu-Adjaye, J. et al. (2015) “An Ecomodernist Manifesto”. Accessed 19 January 2023, http://www.ecomodernism.org/manifesto
(6) Nobel prize Paul Crutzen with Eugene Stoermer have greatly contributed to popularize the Anthropocene concept within scientific community with their eponymous publication « The Anthropocene » in Global Change Newsletter, may 2000, vol. 41, p.17-18. http://www.igbp.net/download/18.316f18321323470177580001401/1376383088452/NL41.pdf
(7) Bonneuil, C. (2022). Comment vivre l’Anthropocène ? EcoRev’, 52(1), 48 68. https://doi.org/10.3917/ecorev.052.0048

Session 9E: The Critique of Technology by French Environmentalist Engineers: a Shift in Values

Antoine Bouzin

Since the end of the 2010s in France, there has been a growing commitment by engineers to ecological issues. This increase in commitment is reflected in the creation of numerous militant organisations at the initiative of engineers and the multiplication of collective actions carried out (Bouzin, 2023). While this ecological commitment is necessarily plural and heterogeneous, some unusual criticisms can be identified in the activist discourse of engineers. Indeed, technologies are regularly denounced as being responsible for the ecological crisis and are sometimes even described as "cogs in a complex set of dominations that are vital to the functioning of the capitalist and extractivist system that causes ecological devastation and sustains social inequalities" (leaflet distributed by the environmentalist association Les Désert'heureuses created by engineers).
This paper thus seeks to analyse the accusations made against technologies by French engineers committed to the environmental cause. The aim is to understand the transformations in the meanings attributed to technologies and the resulting changes in values. Indeed, these criticisms of technologies appear to be in contradiction with the positions and representations traditionally attributed to engineers. Engineers are indeed commonly perceived as defenders of technical progress and a modern conception of technology defined as neutral, objective and rational (Latour, 1993).

On the one hand, we used a sociological method based on biographical accounts to address our issue. We conducted narrative interviews with about fifty French engineers involved in militant organisations dedicated to the environmental cause. We sought to vary several social characteristics in our selection of interviews: gender, age, field of specialisation and business sector. We also used the sociological theory and political philosophy developed by Luc Boltanski and Laurent Thévenot (2006) to understand the situations of conflict between engineers on the issue of technology and to examine the changes in values at work.

This paper aims to present two categories of results:
- Situations from which 'critical moments' are created (Boltanski & Thévenot, 1999). The aim is to identify specific situations in everyday life in which environmental engineers have doubts about the values of progress and common good generally associated with technology.
- Conflicts over which values should be assigned to technologies in the context of ecological crisis. The aim here is to distinguish between the many criticisms levelled at technologies by ecological engineers, to list the various qualifications used and to classify the registers of values expressed.

We observe 'frame disputes' (Benford, 1993) among engineers who are activists for the ecological cause. Technologies are at the heart of disputes in which the characteristic values of the industrial world, performance and efficiency, are vigorously contested in favour of other values of a democratic and environmental nature. This change in the values accorded to technologies challenges the professional practices of these engineers and leads them to question and reorient their technical activities, for example towards low-tech (Bihouix, 2020).

Reference
Benford, R.D., 1993, “Frame Disputes within the Nuclear Disarmament Movement”, Social Forces, vol. 71, n°3, pp. 677-701.
Bihouix, P., 2020, The Age of Low Tech. Bristol, Bristol University Press.
Boltanski, L., & Thévenot, L., 1999, “The Sociology of Critical Capacity”, European Journal of Social Theory, n°3, vol. 2, pp. 359-377.
Boltanski, L., & Thévenot, L., 2006, On Justification: Economies of Worth, Princeton, Princeton University Press.
Bouzin Antoine, 2023, “The language of ecology used by militant engineers”, Socio, n°17, pp. 139-160.
Latour, Bruno, 1993, We Have Never Been Modern, Cambridge MA, Harvard University Press.

Session 9E: The role of “honor” in engineering a solution to climate change

Milenko Budimir

Climate change is arguably the single most important issue of our age, with implications not only for human life and civilization but for all life on the planet for ages to come. Specifically, anthropogenic climate change is caused in large part by human industrial activity beginning with the industrial revolution. This ties it to the work of generations of scientists, engineers, and technologists, making them at the very least partly responsible for the current situation. Singling out engineers (or engineering activity), they bear some responsibility for actions in the past but more importantly for actions taking place today that continue to contribute to climate change. Recently, there have been calls for a Declaration of Helsinki for the engineering profession as a way to address climate change and environmental stewardship (Lawlor and Morley, 2017). Such an action would take seriously the injunctions for environmental stewardship found in many professional engineering societies codes of ethics by going a step further to address climate change, and may work together with existing legal and regulatory frameworks. Still, the question arises, is this enough? Is this sufficient to ensure the type of change that will be real, impactful, and practical? So far to date, with ever more alarming consequences of climate change documented regularly, the answer appears to be no. One possible solution I would like to consider is to explore some work on the concept of moral change. Specifically, Kwame Anthony Appiah in his book “The Honor Code” argues that in historic cases of revolutionary moral change, the mechanism at work was not solely moral argumentation nor legal changes, but rather changes in perception within a group identity about what was considered an honorable action or not. I will explore how a concept of honor appropriate to the engineering profession may help to bring about positive action in combating climate change. I will also consider how some recent examples of disinvestment in fossil fuels can be understood as a change in the types of investments viewed as honorable or dishonorable, and how this change in attitude may serve as a model for engineering as well.

Friday

Session 10A: What does a “Science” of Existential Risk look like? The future of Existential Risk Studies

Gideon Futerman

Existential Risk Studies is a nascent, if growing, area of study, which has particular relevance to those interested in the impact of emerging technology and socio-technical systems on society. The field, thus far however, has lacked an underlying ‘philosophy of science’, a sense of how the field ought to function, what ‘evidence’ in this discipline is, how to be socially organised, what assumptions different approaches are based on and what they mean in practice, and how to inform policy. Over the past few months, we have been exploring how to make a science of existential risk studies, with a particular focus on methodological pluralism and the epistemic utility of differing approaches. This presentation will explore this work, highlighting our vision for the discipline, explaining how evidence is constructed and how and why methodological pluralism is necessary.

We will also focus on the importance of other disciplines in this interdisciplinary field, in particular those working on STS/Philosophy and Ethics of technology, as well as discussing the relevance of Existential Risk concerns to practitioners and theorists interested in emerging socio-technical systems. Some real world applications of some of the different methodologies explored under different approaches will be discussed, which will highlight how we can study specific socio-technical systems in light of their impacts on Existential Risk, and how this can inform action to reduce Existential Risk. Finally, we will discuss useful coalitions we think it is important for Existential Risk Studies to build to help achieve our epistemic and normative aims, and how fruitful collaboration with people interested in Philosophy and Ethics of Technology is vital for the success of the discipline.

Session 10A: The lock-and-key model: Opportunities and challenges for technosolutionism with regard to artificial intelligence

Frieder Bögner

It is often said that technological advances have caused many global challenges that society faces today, most notably climate change, but that many technological developments also suggest promising strategies to address these challenges. There are multiple ways to question this constellation. One of these is the objection that if technology has caused major challenges, it cannot be put into place to solve these problems. This is an attractive approach, but it is far too general as such. It can become more tenable if illustrated with regard to a specific technology sector e.g., replacing the combustion engine with engines that are powered by renewable energy. However, there is a different kind of critique that is advanced against the idea that technologies might lead to promising strategies to combat global challenges like climate change. This is the technosolutionism critique.

According to this critique it is wrong to offer technologies as a solution to societal challenges like climate change since technology is not the appropriate strategy to settle on. So, here the core of the critique does not have to be the idea that the aspect that caused the problem cannot be part of a solution approach. The critique is potentially more far reaching instead since it can take other reasons as its foundation. Looking at recent developments it is especially the attitude and strategy to use artificial intelligence applications as solutions to climate challenges or to mitigate their causes which becomes the target of a technosolutionism critique. What gets criticised is the idea of putting innovative and high-performance technology into the service of a comprehensive societal challenge. This can be part of the technosolutionism critique.

It is the aim of this talk to propose a more refined reading of this critique towards technologybased solutions, suggest options how to react to this objection and apply these findings to the case of artificial intelligence that is employed to combat climate change related challenges or obstacles to sustainability in general. Technosolutionism is an important kind of criticism since the topics of technology, societal challenges and solutions strategies get connected with regard to the discursive space involved in this debate. That is, by connecting these issues, on can reach a better understanding of the argumentative options available around that particular kind of criticism.
In the talk, I will first address the question what motivates or explains the resort to technological approaches to address societal challenges. Second, it will be illustrated which kind of problems and solutions strategies are relevant for the debate around technosolutionism. In the third step, a particular refined understanding of the technosolutionism critique is laid out, the lock-and-key model of technosolutionism. In a final step I will draw to some artificial intelligence applications and evaluate in how far these fit to the proposed lock-and-key model. Overall, this talk deepens the understanding of a particular aspect of the debate around artificial intelligence in service for societal challenges and the attitudes involved.

Session 10B: Understanding Designing to Match Users’ Mental Models as Eidetic Reduction

Dylan Wittkower

Technology that doesn’t work the way that the user expects it to is confusing, difficult to use, and sometimes dangerous, so an important part of engineering and design practice is understanding the user’s mental model of the technology in order to match it as closely as possible and to anticipate and mitigate harm that may follow from any ineliminable divergences. It has not been sufficiently explored that the process of discerning the user’s mental model, central to technology design, is fundamentally a phenomenological project in the sense specific to philosophy. Once this connection is made, a wide variety of valuable tools and methods from phenomenology and post-phenomenology become ready-to-hand for engineering practice.
That this aspect of engineering practice is fundamentally a phenomenological project can be demonstrated by describing the process using Husserlian terminology. The user’s “natural attitude” (naturliche Einstellung) to the technology is realist; they expect the reality of the technology to correspond to the mental model (“representation,” Vorstellung) that the user intuitively and uncritically forms of it. The user’s mental model/Vorstellung is always already conditioned by and formed through their intersectional lived experience and history and within the context of use; these user-relative elements of the mental model/Vorstellung are not separated in the user’s experience from its objective attributes—this is the way in which the mental model appears through Givenness (Gegebenheit) within lived experiences (Erlebnisse). To articulate the mental model, the engineer or designer has to temporarily ignore the actual structure of the technology as well as design constraints and priorities falling outside of the user’s mental model. For example, it may be obvious to a designer that a wifi-dependent IoT device will cease to function properly when removed from its place in the home of office, but the designer must “bracket” (einklammern) this knowledge in order to focus on whether this will be given within the user’s lived experience. This phenomenological reduction or epoché to the user’s lived experience is necessary to describe the user mental model, but the description itself also requires an eidetic reduction. Imaginative variation (Phantasievariation) is then used to determine the user mental model and how it arises from different aspects of how the technology is given within their lived experience.

To demonstrate the value of understanding technology design as including a phenomenology of the user, we can consider at least a couple examples of how tools and methods from phenomenology and post-phenomenology can be applied to engineering practice through this connection. In the time remaining following the primary argument above, I will provide two examples. The first uses Heideggerian deconstruction (Destruktion; Abbau) to explain why people from Anglo-European cultures are much less comfortable with robots than those from other cultures, for example those of East Asia. The second uses foregrounding and backgrounding in the phenomenology of perception to explain the user behavior that gives rise to the so-called “privacy paradox.”

Session 10B: Operator guidance systems in road construction: a technological mediation perspective

Hans Voordijk, Seirgei Miller &
Faridaddin Vahdatikhaki

The quality of roads depends heavily on the proper compaction of the asphalt layer. Compaction is an intricate operation that is influenced by a myriad factors: design (e.g., the type of asphalt mix), execution (e.g., uniformity of compaction), logistics (e.g., the temperature of the asphalt at the delivery time), and environment (e.g., weather condition). In practice, operators rely on their experience to cope with the dynamic aspects of compaction. An outcome of this dependency on operators' tacit knowledge and skills is the high variability in the quality of the compaction operation, and thus the asphalt layer.

Existing support systems collect both design (e.g., pavement design) and real-time (e.g., asphalt temperature and roller pass count) data to help operators better plan (or re-plan) and execute compaction operations. The cognitive load imposed by these systems can be reduced by transferring the task of data interpretation and integration to the system as much as possible. Operator support systems can be transformed from descriptive systems, i.e., systems that only provide information needed for planning a compaction operation, into prescriptive or operator guidance systems, i.e., systems that suggest actionable compaction strategies based on the collected data.

The ways in which operator guidance systems in road construction translate data into actionable guidance for operators can be analysed from the technological mediation approach (Ihde 2009). From this perspective, operator guidance systems can be understood as mediators between their users and the roads to be paved. The objective of this study is to understand the mediating role that these systems play between their users and characteristics of the roads that are in a great part unperceivable by the operators.

Major question is how perceptions and actions of users are mediated by operator guidance systems. In answering this question, the well-known typology on human technology relations of Ihde (1990) is used first. Each relationship addresses a specific way that technologies mediate human perceptions of reality. Second, the framework developed by Dorrestijn (2017) is used. This framework translates the technical mediation approach to practice and provides more detailed insights in the ways how operator guidance systems influence or transform users' actions.

A case study of a specific application of an operator guidance system supports the philosophical reflection on technological mediation in a road pavement project. Technological mediation through such a system makes certain properties of roads to be paved accessible to operators. In the case study, detailed information is provided about the process of how 'input' in the form of data from cameras and sensors is transformed into actionable 'output'. Operators turn to these systems in order to interpret the guidance they provide (Rosenberger and Verbeek 2015). The scientific contribution of this study is in connecting the operation of these systems to this technological mediation perspective.

Session 10B: Operator guidance systems in road construction: a technological mediation perspective

Andrej Dameski, Andreas Spahn, Alessandro Corbetta, Antal Haans, Caspar A. S. Pouw, Federico Toschi, Jaap Ham, Rabia I. Kodapanakkal &
Gunter Bombaerts

Although not without its challenges, the study of human social behaviour using empirical data is not untypical. At train stations in particular, empirical data is commonly used to study movements of pedestrians, formation of jams and crowds, efficiency of boarding and deboarding, train car seat allocation and occupation, use of kiosks and other auxiliary areas of the platforms, freeriding, and other phenomena. However, most of the social, ethical, and philosophical interpretation of this type of empirical data is focused on individuals or individual-inspired explanations.

During our research project named SRCrowd (socially responsible crowd management) we have opted to focus on collectives, such as crowds and groups. We explore how they affect the design, employment, and the (ab)use of technology, with the help of empirical data and analysis provided by the sister project of HTCrowd (a high-tech platform for human crowd flows monitoring, modelling, and nudging). Our overarching hypotheses are that a) collectives do have a distinct and substantial effect on technology and on individual users; b) individuals and technology also influence collective formation and behaviour; and that c) (some of) these effects can be studied using empirical data from Dutch train stations.

To give structure to our research output as well as to improve our power of interpretation, we have opted to take the theory of technological mediation (itself arising from postphenomenology, and part of philosophy of technology), and complement it with a collective perspective and systems theory. For this purpose, we suggest thinking of mediation as a triangle between individuals, collectives, and technology, rather than as a relation solely between humans and technology. We then use this theoretical framework to study and interpret the empirical findings focused on the social behaviour of train platform visitors.

Our findings so far indicate that from our empirical data we can confirm that collectives do have a distinct and substantial effect on how train station technology is designed, employed, and used. This also includes their mediative effect on how individuals use technology. The effect applies both ways, thus technology and individuals also substantially affect collectives (besides affecting themselves). For example, it was found that platform design affects crowd formation i.e. boarding crowds closer to stairs and escalators are more homogenous in size and population than those further away. Crowds also affect individual travellers by mediating their use of trains as technology, through compelling them to wait in queue before boarding or deboarding. It was also found that boarding and deboarding crowds interact with 'push' and 'pushback' mechanisms, which result in a particular deboarding corridor width and shape. We argue that these empirically attested phenomena are difficult — if impossible — to be explained by using a purely or predominantly individualistic explanations, and warrant collective interpretation as a supplement.

Thus, our provisional conclusion is that adding collectives to the focus study together with individuals can provide an enhanced picture of human behaviour around technology, as well as provide explanation for some phenomena that are difficult or cumbersome to tackle from a predominantly individualistic perspective.

Session 10C: Technology-mediated moral change. Exploring the relation between moral concept change (MCC) and moral value change (MVC)

Samuela Marchiori

Technology has the potential to profoundly impact moral systems. This can lead to moral change, i.e., the "systemic pattern of behavioural and cognitive changes [...] that is observed across a sufficiently wide population [and] commonly described using moralised language and evaluated in terms of moral emotions" (Danaher & Sætra, 2022, 3). In this paper, I refer to technology-mediated moral change to highlight the interconnectedness between technology and society, whereby one cannot be properly investigated in isolation from the other (Verbeek, 2011), as well as to account for types of moral change that may not be induced by technology (van de Poel, 2021).

The focus of the paper is on the relation between technology-mediated moral value change (MVC) and moral concept change (MCC), which remains largely underexplored, despite a growing body of literature on both dimensions of moral change (Boenink & Kudina, 2020; Hopster, 2022; Löhr, 2022; van de Poel, 2022; van de Poel & Kudina, 2022; Veluwenkamp et al., 2022). Specifically, the literature suggests that certain types of MVC are instances of MCC; it also suggests that MVC can occur even when MCC does not (van de Poel, 2021). However, the relation between the two has not been explored further. This raises the question of whether MCC is a subcategory of MVC, or whether the two are merely overlapping. That is to say, is every instance of moral concept change also an instance of moral value change?

In this paper, I attempt to answer this question by clarifying the distinctive features of technology-mediated MCC and MVC and presenting a tentative taxonomy thereof. Based on this taxonomy, I will argue that, while certain types of MCC may fully overlap with instances of MVC (e.g., when thin moral concepts are concerned), that is not always the case (e.g., when thick moral concepts are concerned). Therefore, moral concept change is not a subcategory of moral value change. Lastly, I note that the relation between MCC and MVC largely depends on the level of abstraction at which both concepts and values are understood (van de Poel, 2021).

References

Boenink, M., & Kudina, O. (2020). Values in responsible research and innovation: From entities to practices. Journal of Responsible Innovation, 7(3), 450-470.
Danaher, J., & Sætra, H. S. (2022). Technology and moral change: The transformation of truth and trust. Ethics and Information Technology, 24(3), 35.
Hopster, J. (2022). Future value change. Prometheus, 38(1), 113-123.
Löhr, G. (2022). Do socially disruptive technologies really change our concepts or just our conceptions? Technology in Society, 72, 102160.
van de Poel, I. (2021). Design for value change. Ethics and Information Technology, 23, 27-31
van de Poel, I. (2022). Understanding value change. Prometheus, 38(1).
van de Poel, I., & Kudina, O. (2022). Understanding Technology-Induced Value Change: A Pragmatist Proposal. Philosophy & Technology, 35, 40.
Verbeek, P. P. (2011), Moralizing Technology: Understanding and Designing the Morality of Things. Chicago: University of Chicago Press.
Veluwenkamp, H., Capasso, M., Maas, J., & Marin, L. (2022). Technology as driver for morally motivated conceptual engineering. Philosophy & Technology, 35, 71.

Session 10C: The Role of Artifacts in Value-Formation: From Affordance to Exposure

Aznavur Dustmamatov

In recent years, the concept of affordance was enlisted to explain how artifacts can embody or promote certain values (Klenk; Tollon). An artifact is said to promote an activity by providing, through its features and design, a stronger “force” of affordance (Tollon 2022), e.g. a car with its internal combustion engine affords speedier transportation than a horse carriage.

I argue that, in this articulation, the notion of ‘affordance’ can only explain how artifacts amplify the expression of pre-existing values, but it cannot explain how artifacts contribute to value-formation. This is because 'affordance' presupposes an average, idealized use-context that is independent from the presence of the artifact itself. This supposition precludes the possibility of understanding the reflexivity exhibited by artifacts: their exploitation at scale transforms the very context in which they are embedded.

I reconceptualize affordances as something that artifacts possess only as embedded in the definite context of a “sociotechnical system” (van de Poel). An artifact’s affordance is relative to a context of instrumentation (e.g., in some terrains a car might be slow), while the “force” of this affordance is relative to a context of evaluation that defines the value of a given outcome.

First, I consider how artifacts, when made available at scale, transform their contexts of instrumentation. Artifacts generate and sustain new sociotechnical systems such as the upstream ecosystems of production and distribution, and the downstream ecosystems that rely on exploitation of the artifact. These ecosystems can in a circular fashion affect the original affordance, such as when the proliferation of drive-thrus and suburban sprawl in some areas transformed the car from a speedier alternative to a basic necessity of life.

Second, I consider how artifacts transform their contexts of evaluation. Evaluation relies on a set of criteria for selecting actions from among the afforded possibilities. The selection is based on the differentials that are salient to a specific possibility-space. Artifacts enable new possibilities or alter the cost-reward trade-offs within a given possibility-space. Changing the make-up of the possibility-space may lead to a shift in salient differentials: for example, if none of the previous possibilities had feature F, then F could not serve as a differential, yet it could become a differential if a new possibility with feature F were introduced. Since evaluation ranges over the exposed possibilities only, values may not change until full exposure to the possibility takes place, e.g., although asbestos had been toxic all along, there was no change in policy or attitude until its hidden costs were revealed to the public. By introducing new exposures, artifacts can shift differentials.

Session 10C: The Dynamics of Technological Value-Ladenness: A Pragmatic Approach

Byron Newberry

The belief that technologies are value-laden has a long pedigree in the scholarly literature of philosophy of technology and technology studies, as well as in popular writing about technology. Value-ladenness is taken to mean that human values are somehow embedded in, embodied by, or transmitted through technological artifacts. This is as opposed to a value-neutral view of technology, perhaps held by a minority of philosophers, and often assumed to be prevalent amongst the practitioners of science and technology. The value-neutral view rejects that technological artifacts can embody any values beyond the instrumental value of usefulness as tools.

While the notion of technological value-ladenness enjoys considerable support in the scholarly literature, proponents of value-ladenness nonetheless have numerous non-trivial questions to answer in defending the idea. What is the precise mechanism for value-embodiment? How does one recognize embodied values in a technological artifact? How do embodied values influence users of artifacts? Are the values embodied by an artifact fixed and static, or can they change? If they change, how does that change happen?

This presentation addresses the latter questions about value change in technological artifacts. In prior work [1], I have articulated a pragmatic account of value-ladenness that draws upon the notion of values as relations from John Dewey, as well as the notions of "goodness in a way" from Judith Jarvis Thomson and of embodied values as affordances from Klenk [2]. Here, I will use my pragmatic account to address the dynamic nature of embodied values. Consistent with the pragmatic account of technological values given by van de Poel and Kudina [3], embodied values on my account can be idiosyncratic (particular to an individual or to a circumstance) or shared (prevalent across many individuals and circumstances). The latter case of shared values is particularly important since it implicates potential societal-wide effects of technology.

Further, using the value-as-relation perspective I argue that the nature of the values embodied by a technological artifact will depend upon the nature of the relationship between valuer and artifact. That relationship will change given changes in either the artifact, the valuer, or both. This may cause new embodied values arise, current embodied values to evolve, or previous embodied values to diminish. I will illustrate this with examples involving malfunction, obsolescence, and innovation. I will also highlight the notion of non-use value, which I take to mean embodied values that are dependent upon artifact functionality, but not upon artifact use. I will show how such embodied non-use values can fluctuate with time.

References
[1] Redacted for review
[2] Klenk, M. How Do Technological Artefacts Embody Moral Values?. Philos. Technol. 34, 525–544 (2021). https://doi.org/10.1007/s13347-020-00401-y
[3] van de Poel, I., Kudina, O. Understanding Technology-Induced Value Change: a Pragmatist Proposal. Philos. Technol. 35, 40 (2022). https://doi.org/10.1007/s13347-022-00520-8

Session 10D: Maintenance & Philosophy of Engineering

Mark Thomas Young, Rebecca Mossop, Alexandra Karakas & Steffen Steinert

Throughout the course of their professional lives, engineers do more than just engage in the design of new artifacts. According to some estimates, the work of the majority of engineers concerns existing artifacts, either through their inspection and the development of methods for controlling the changes they undergo or through the planning of retrofitting operations whereby they are modified to align with new needs, interests and values. In other words, engineers do not only create – they also maintain. Yet despite a recent surge in interest in practices of maintenance among scholars in STS, anthropology, history of technology and geography, maintenance still remains a neglected topic in the philosophy of engineering. This is all the more surprising when we consider the extent to which the examination of maintenance reveals a range of philosophical issues which remain unaddressed by existing literature on the topic. The goal of this panel is to address this neglect, by exploring philosophical issues which arise from considering the different ways in which engineers engage with artifacts after they're built or produced. The panel consists of four individual presentations, each of which highlight different philosophical issues arising from the practice of maintenance by engineers and demonstrate the rich potential of the topic of maintenance to the philosophy of engineering. In addition to exploring complexities surrounding the meaning of the concept of maintenance itself, the presentations in this panel also illustrate how attending to technologies across time reveals new ways of conceptualizing the relationship between technology and values, malfunction and the nature of engineering practice.

Evolving Infrastructure: Maintenance and the Political Lives of Artifacts

Mark Thomas Young (University of Vienna)

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 activity of maintenance. In this presentation my goal will be to demonstrate how the neglect of maintenance 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. In order to illustrate how attending to maintenance reveals the inherently temporal nature of the politics of artifacts, I'll focus upon the maintenance history of a well-known artifact: the Golden Gate Bridge in San Francisco. My goal will be to illustrate how the maintenance history of the bridge is also at the same time a political history. In order to do so, I'll detail how since construction was completed in 1937, the bridge has undergone a wide variety of different changes in response to the changing social environment in which the bridge exists. After detailing reasons why maintenance should be understood as the central mechanism through which such changes occur, 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 the bridge requires alteration in the form of repair or rehabilitation, 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 these processes are revealed in the maintenance histories of artifacts by charting the political debates underlying the transformations of the Golden Gate Bridge.

Communicating smoothly: how material breakdown is not the only thing to be maintained in a Large Technological System

Rebecca Mossop (University of Luxembourg)

In this presentation, I'll explore the example of the Luxembourg telephone system (1885-1990) to show how different types of maintenance are and were fundamental to this essential infrastructure and its development in a city and global context. Whereas existing discussions of the maintenance of large technical systems (LTS) have focused mainly on material breakdowns and ad-hoc repairs, I'll elaborate a different conception of maintenance as an activity carried out by users that puts emphasis on the matter and its master. While within a LTS this alone indicates the dependency on the master (state, authority, company, etc.) and therefore creates questions of norms of groups, laws and leading technologies, in a communication LTS, I will show how it stimulates negotiations about what is to be maintained, if at all, and how this is to be achieved. Here we see how maintenance also becomes a question of speech acts (Searle 2017) in technological stages of upheaval. Moving away from breakdown as a material concept and maintenance as a question of economic necessity by institutions, my analysis reveals instead how deficiencies in the system (i.e. too few connections in a telecommunication system, not the latest technology, new global standards in the industry or disruptive economic ideas to expand usage) become a declared problems to be fixed, leading to perceptions of the old system as deficient. Such declarative speech acts possess the capacity to change whole systems according to the will of the group which is collectively accepted. In communication systems, these are cases in which a functioning system is described as no longer up to date, deficient and new strategies are tested and developed to maintain the new collective imagination of the system.

How can Science and Technology studies benefit from the Philosophy of Maintenance?

Alexandra Karakas (Budapest University of Technology and Economics)

In technology and engineering, maintenance means many things: repairing, conserving, preserving, reconstructing, and replenishing, to name but a few. These different activities generally aim to keep an artefact functional and prevent or restore malfunction. Thus, maintenance centres around the idea of actual or possible failure. However, at first sight, failure – especially from a user's perspective - seems to be a simple concept, but in reality, both in the philosophy of technology and engineering failure as such is a wide-ranging phenomenon. Even though in engineering, failure is a central issue, and there is a significant amount of research about failure. Still, there is no agreement on the exact meaning of failure or malfunction. A common feature of the definitions of failure is that they are binary. The talk focuses on how a binary notion of function and malfunction cannot, on the one hand, cope with the array of artefact performances and, on the other hand, how it cannot fit into an extensive maintenance philosophy either. I propose a gradual notion of malfunction, which is more adapted to different maintenance strategies and better explains the subtle ways artefacts work in engineering and science. In a case study about the OPERA experiment, I analyse issues connecting to the philosophy of science rooted in the problem of malfunction and maintenance. The case study also shows how science and technology studies can benefit from a maintenance framework and reveal theoretical consequences of faulty engineering.

Making things durable: Maintenance of technology as maintenance of society and values

Steffen Steinert (Delft University of Technology)

In the presentation, I will make the case that thinking about maintenance should look beyond technical functions and take technology's social and value aspects seriously. I will explore how technology maintenance is related to social and value maintenance. Specifically, I will focus on how technology use can intentionally and unintentionally reproduce values and social structures. My goal is to provide some theoretical tools for making sense of how maintenance of technology, the maintenance of society, and maintenance of value hang together. To lay the groundwork, I will make plausible that the notion of maintenance can include both the technical and the social domain. After that, I will make the case that maintaining technology is at the same time maintenance of social order and values. After that, I will focus on the maintenance of socio-technical systems. This is necessary because most technology comes in the shape of socio-technical systems, which are intricately linked to society. Finally, I will turn to the relationship between the maintenance of technology and the maintenance of value and provide an account of how we can make sense of the maintenance of value.

Session 10E: Value Change and Engineering

Ibo van de Poel & Freek van der Weij

1. Assignment

This workshop is intended to familiarize participants with different methods for dealing with value change in design and engineering. In the white paper of the ERC project 'Design for Changing Values', shared alongside this document, a variety of these methods are described. After the introductory talk by Ibo van de Poel, we ask you to form groups of 3-5 people. With your group, you will be asked to make a poster describing how one or more of the methods for dealing with value change should be applied to the development of **self-driving cars**.

Step 1: Select and develop a scenario where value change might occur (see section 4 for ideas)
Step 2: Propose how one or more of the methods mentioned in the white paper should be used to account for/analyze this type of value change. The goal here is to develop a proposal that could be executed in practice. It should not be overly ambitious, and choices for selecting a method should be justified. We have a panel of five experts who are here to answer your questions regarding a particular method for dealing with value change, see below.
Step 3: Provide proposals for design requirements based on your investigations.

2. Expert panel

Value monitoring	Tristan de Wildt
Technomoral scenarios	Olya Kudina
Prototyping	Joost Mollen
Simulation methods	Anna Melnyk
Design principles	Ibo van de Poel

3. Time schedule

09:00 - 09:20	Introduction
09:20 - 09:25	Group formation
09:25 - 10:10	Exercise
10:10 - 10:30	Feedback on posters

Session 11A: Art-Inspired Pedagogies in Engineering Education - Using Comics, VR/AR, Gaming, and Music in Engineering Education

Kai Zhuang, Dimpho Radebe & Mojgan Jadidi

In recent years, there is a growing recognition in engineering education that creative, humanistic, and transferable skills such as emotional intelligence, ethical leadership, and teamwork, are essential to students' success, thriving, and contribution in university and beyond. However, most engineering students are used to rigorous curriculums that emphasize technical development, with little opportunity to explore creative and humanistic subjects and to develop ethically and holistically. Moreover, many engineering students who are used to highly reductionist and analytical thinking find it difficult to engage with "softer" learnings, and may experience lowered motivation in these subjects.

Over the past three years, partially in response to the pandemic, we have been exploring a number of art-inspired pedagogies, sometimes with support of emerging technologies, to support engineering students' learning journeys. Specifically, we have explored the use of comics, virtual reality (VR), augmented reality (AR), computer games, and music in the classrooms. We have applied these learnings to both technical and humanistic learnings.

Through these early stage pilots, we noticed three benefit of these art-inspired approaches:

[1] Our emphasis on experiential learning and storytelling humanized the content, making them more engaging and appealing for students.
[2] By combining visual and verbal elements through storytelling and interactive games, we help students engage both the analytical left brain and holistic right brain, making it easier for them to grasp complex ideas and promote creative and associative thinking.
[3] By making use of media formats (comics, 3D gaming, AR/VR) that many engineering students are either familiar with or interested in, their motivation in learning is improved.
[4] Although words are important as they are essential for conceptualizing and communicating information, we also recognize that words can be pedantic and inaccessible. When considering the diversity of backgrounds present in the North American engineering students' body, and that English may not necessarily be every student's first language, art-inspired pedagogies can present a lowered entry barrier for new information.
[5] The use of these new pedagogies, particularly with the use of VR and AR, can replace some laboratory equipment needs, making engineering learning more accessible.

It is our hope that by continue our exploration of these art-inspired pedagogies, we can forge a new pathway for developing engineering students holistically and readying them for the emerging future.

Session 11A: Situating values in practice: University teachers' perspectives on assessment and emerging AI

Elin Sporrong, Cormac McGrath & Teresa Cerratto Pargman

AI in education (AIED) is proposed to increase efficiency of assessment, support individualization (Swiecki et al., 2022) and reduce teachers' workload (Chaudhry, 2022) by the use of machine learning in automated grading, assessment and exam proctoring (Gardner et al., 2021). Anticipations of AI in higher education (HE) motivate changes of policies, mindsets and allocation of resources (Bechtold et al., 2017) as well as an abundance of ethical guidelines for AI-development and usage (Dignum 2019).

However, broad guiding principles and abstract values, provide little insight about educational stakeholders' needs and value-interpretations, which may be critical in order to address tensions between current and potential practices (van der Velden & Mörtberg, 2014) and possible negative effects of technology (Holter, 2022).

Research in education and technology studies consistently shows that teachers' perspectives are missing in AI design and decision making (Barton, et al., 2020), leaving great discrepancy between the potentials of AIED and their implementations in education (Zhang & Aslan, 2021).

This interview study aims to tackle this knowledge gap, by exploring: i- what teachers find important and valuable in their current assessment practices and, ii- teachers' perspectives on how these values could be amplified or reduced (Rosenberger & Verbeek, 2015) connected to emerging AIED in assessment practices.

To collect data, we targeted university teachers from one of the main universities in Sweden. Group and single semi-open interviews were held with eight HE-teachers from the faculty of law, natural and social sciences and the humanities. To formulate a transparent method for value-elicitation (Dignum, 2022), the interview questions and interview prompts were drawn upon Boenink & Kudina's (2020) practice-based approach to values, understood as interactive, dynamic and lived realities.

In progress analysis of the collected data (i.e. interviews transcribed verbatim), builds on a post-phenomenological approach (Rosenberger & Verbeek, 2015). The aim of the study is to contribute with critical insights about teachers' perspectives on potentials of AI and values that risk being lost (Selwyn, et al., 2021) in practice.

Preliminary findings highlight that teachers' valuing processes vary connected to e.g academic discipline, assessment tasks and values. Broad value concepts, such as transparency, can vary by context, and connect with e.g.: i- a safe working environment, if a lack of transparency could increase the risk of harm, ii- justice connected to assessment or iii- legitimacy in assessment design and communication with students.

Emerging AIED, can potentially amplify or reduce perceived values in teachers' practices, depending on e.g. how adjustable the technologies' functionalities are and how well they align with practical preferences and moral values, as these are interconnected in teachers' practices in complex ways.

Approaches such as Value sensitive design (Friedman & David, 2019) that mainly consider broad human values in technological development, and that do not sufficiently consider the situated valuing processes of stakeholders, might thus not fulfill the intended aims of value alignment for AI-driven assessment technologies.

Session 11A: Rethinking engineering education in a French-speaking context using the concept of Bildung

Christelle Didier

Today, the German concept of Bildung has become visible as a threshold concept in science and environmental studies where critical thinking, societal values, ethics, and social responsibility have been introduced as new knowledge domains (Sjöström et al. 2017). In this paper we set out to discuss whether rethinking engineering professionalism through the concept of Bildung would be beneficial for addressing technological issues as well as broader socially pertinent challenges on a global scale such as global warming, pandemics, decreasing biodiversity, increasing economic and social polarization, racism, information- and cybercrimes, etc. by which society is presently confronted. Dealing with such an array of challenges will have bearings on understandings of the very telos of engineering (Aristotle 2000, Buch 2022). Rethinking the telos of engineering has profound implications. Ultimately it is a struggle over knowledge as such endeavour will do away with the technical core/nontechnical periphery distinction in traditional engineering epistemology.

The paper starts with a reflection on the present configuration of higher education and the priority given to STEM fields. The concept of Bildung and understandings of Bildung usually rise to prominence when a given configuration of the educational system is no longer able to deliver on its promises (Christensen 2019). Bildung can therefore be seen as a counter and problematizing concept vis-à-vis the present configuration of the educational system.

As regards engineering, it is a heterogeneous socio-technical enterprise in which challenges and solutions are distributed across social, material, economic, political, and technical domains. Various understandings of "Bildung in Engineering Education" have been identified: 1. Wilhelm von Humboldt´s classical Bildung, 2. Anglo-American liberal education, 3. Scandinavian folk-Bildung, 4. Democratic education, and 5. Critical-hermeneutical Bildung (Sjöström et al. 2017). Considering the negotiated nature of engineering solutions, together with the heterogeneous character of its practices, we argued that option 4 has specific relevance. The paper thus points to John Dewey's ideas of democratic education and unfolds elements in Dewey's philosophy that are helpful for a contemporary reconstruction of Bildung as democratic education (Dewey 2011).

In so doing it is necessary to address the very purpose of engineering – its telos. Considering the challenges facing humankind in the Anthropocene era, this telos will have to be reconsidered. We use Dewey's pragmatist lens to do so and argue that Engineering should be reconsidered as a problem-solving discipline that seeks to ameliorate living conditions. In the conclusion, we consider possible obstacles for reconstructing engineering along the lines of Bildung through democratic education.

References

Aristotle (2009). The Nicomachean Ethics, OUP Oxford.
Buch, Anders (2022). The Need for Recovery of Engineering. In: Christensen, Steen Hyldgaard et al. (eds.) Engineering, Social Sciences, and the Humanities: Have their Conversations Come of Age? Springer Nature.
Christensen, Steen Hyldgaard (2019). Costs and Benefits of Commercializing Teaching, Research, and Service in the American Corporatized University. In: Christensen, Steen Hyldgaard et al (eds.) The Engineering-Business Nexus: Symbiosis, Tension and Co-evolution. Springer Nature.
Dewey, John. (2011) Democracy and Education, Simon & Brown.
Sjöström et al. (2017). Use of the concept of Bildung in the international science education literature, its potential, and implications for teaching and learning, Studies in Science Education, 53(2).

Session 11B: The Development of the Affordances of Indicator Diagrams into Thermodynamic Work Cycles in the Material and Theoretical Cultures of Early Steam Engines

Alok Srivastava

The thermodynamic work cycle is a central piece of model-based reasoning and diagrammatic reasoning for investigating and designing energy conversion processes in engineering, physics, chemistry and biophysics. The model and diagrams of thermodynamic work cycles are applied to material processes and is a commonly used logic and language in the investigation and design of these processes. In the early 1800s, Sadi Carnot developed this machinery of the thermodynamic work cycle by analyzing the indicator diagrams generated by recording devices mounted on steam driven engines to trace defects in performance of the steam engines. However this episode in scientific discovery and conceptual change has not been studied concurrently for the affordances and semiotics of indicator diagrams and development of model-based reasoning. The affordances of the indicator diagram were developed and articulated by Sadi Carnot's clarification of the semiotics of the traces across many types of engines. The traces of the movement of the piston during a work-cycle had to be abstracted as referring to a conceptual and reference work cycle. This work resulted in the generation of the epistemic objects of the scientific and engineering disciplines of thermodynamics. In this paper, I intend to trace the sequence of steps of development and the evolution of local affordances of indicator diagrams into the semiotic, logic and language of engine design and in turn instantiated and mobilized the metaphysical affordances into the semiotics, logic and language of thermodynamics as developed by Carnot and his successors. The overall goal of this paper is to develop tools for elucidating the specific steps of development of effective ontologies, logics and language from diagrammatic reasoning to explanatory and model-based reasoning during episodes of scientific discovery and conceptual change. (Resource Paper: Wise, M. N. (2018). On the Stories Told by Indicator Diagrams and Carnot Diagrams. Endeavour, 42(2-3), 145-156.)

Session 11B: Christiaan Huygens and the Foundations of the Engineering Thermodynamic Worldview

Terry Bristol

The foundations of the engineering thermodynamics worldview emerged in Europe (1650-1850). These foundational insights remain underappreciated in the modern era. Cardwell realized that "accounts of the concepts of work and energy have tended to describe them within classical mechanics. ... This takes too narrow a view." Mechanical accounts of thermodynamics are, what Feyerabend called, 'rational reconstructions'.

Gillispie characterizes the period by the transition from the science of mechanics to the science of machines. Leibniz proposed a corresponding meta-paradigm shift from Statics to Dynamics. In 1803, engineer Lazare Carnot explicitly proposed an engineering thermodynamics worldview.

But where did Carnot's mature ideas come from? Lazare traces them back through Maupertuis and Leibniz, and, as with all the other key contributors, acknowledges links to Galileo and to the geometry of Archimedes. Other contributors included d'Alembert, the Bernoulli's, Euler, and Newton, to mention a few. My focus here is on Christiaan Huygens.

Huygens investigations led to a more comprehensive framework. I will emphasize two components of Huygens work: his theory of motion and his understanding of the compound pendulum. Both these depend on his Center Frame methodology.

His theory of the elasticity of interactions led to his theory of motion. His Center Frame reasoning allows us to differentiate the Static Galilean relativity and the Dynamic Huygens relativity. Galilean relativity is appropriate to Special Relativity. Huygens relativity, incorporating acceleration, is more appropriate to General Relativity.

Second, Huygens recognized the 'circular error' of Galileo's idealized pendulum motion. He realized that the cycloidal path always involves two, compensating, complementary components. These can initially be thought of in mechanical terms as positional potential (mv) and kinetic potential (mv2). All motion, all change, is composite, with an irreducible component of each. This introduces a new concept of change (motion) requiring a more general representation. What emerged was the concept of action.

Using his Center Frame reasoning Huygens identified the engine-like cyclic properties of the compound pendulum. He realized the need for more comprehensive, post-mechanical engineering concepts of continuity, symmetry and conservation (of MV2).

References

Huygens, C. (1673) Horologium Oscillatorium; (1703) De motu corporum ex percussione
Blackwell, Richard J. (trans.) (1977) "Christiaan Huygens' The Motion of
Colliding Bodies", ISIS, 1977, 68 (No. 244)
Gillispie, Charles (2014) Lazare and Sadi Carnot – A Scientific and Filial Relationship
Carnot, Lazare (1803) Foundational Principles of Equilibrium and Motion; and Geometry of Position
Ducheyne, Steffen (2008) Galileo and Huygens on free fall: Mathematical and methodological differences. Dynamis [online]. 2008, vol.28, ISSN 2340-7948.
Stan, Marius (2016) "Huygens on Inertial Structure and Relativity", Philosophy of Science, 83 (April 2016) pp. 277–298.
Feyerabend, Paul (1981) Problems of Empiricism Philosophical Papers Vol. 2
Cardwell, Donald (1971) From Watt to Clausius: The rise of thermodynamics in the early industrial age
Tho, Tzuchien, (2017) Vis Vim Vi, Declinations of Force in Leibniz's Dynamics
McDonough, J, (2022) A Miracle Creed: The Principle of Optimality in Leibniz's Physics and Philosophy

Session 11C: The individual and Civil Engineering

Dennis Gedge

The tools of the trade of civil engineers, have changed dramatically like many technological things in recent years, but the fundamental job of civil engineering remains the same, and it is recognisable from ancient times. Questions about professionalism and the individual engineer are still with us.

Current professional guidance to engineers is that climate change is the biggest challenge civil engineering faces at the moment, however , it has always been a matter of dealing with the constraints of the laws of nature as well as the laws of the land. Failures in dealing with natural forces have always been obvious fairly soon after construction. But the sinister effects of heavy construction upon the atmosphere and other parts of the natural environment have slowly crept up unnoticed.

The paradox is that providing essential things like road improvements, and maintaining the safety of structures is not optional, it is a job that just has to be done. It cannot be switched off like harmful things that individuals may do to themselves.

Civil engineering is a profession scattered within and industry. Politicians decide which headline projects are good and which of them should be constructed. The public, (and perhaps the professional institutions too) have a romantic view that Civil Engineers are more influential than they really are. The institutions are groups of individuals, not trade organisations for commercial companies.

Current ethical guidance from the professional institutions concentrates on straightforward problems like engineers being coerced to make statements or designs for financial reasons to suit their employers. But the most interesting unanswered questions remain, such as identifying exactly where responsibility might lie, and the relationship between legal and professional duties.

The Civil Engineering profession is signed up to the U N Sustainable Development Goals, but is the idea of politicians working together with a diffuse profession just a dream?

Has an understanding of the changing climate brought any new questions about professionalism for Civil Engineers? Good engineering projects are the efficient ones, the ones that minimise the use of energy, individual engineers do have a hand in that.

Reference
An engineer is someone who can do for a shilling what anyone can do for a pound. Anon.

Session 11C: Practitioners' reflections on organizing Transdisciplinary Innovation

Marc Steen, Jurriaan van Diggelen, Friso van Houdt, Robin van den Akker, & Leanne Cochrane

Our paper deals with organizing Transdisciplinary Innovation (TI). This is an underlying theme of fPET, which brings together philosophy and technology. Moreover, TI is critical in enabling experts in technology to collaborate with experts from disciplines such as ethics, law, or social science.

There is some knowledge on transdisciplinary research (e.g., Bernstein 2015). Rather less is known, however, on transdisciplinary innovation, which McPhee et al. (2018) define as: 'not just about working towards a shared goal or having disciplines interact with and enrich each other. Instead, transdisciplinary innovation is about placing these interactions in an integrated system with a social purpose, resulting in a continuously evolving and adapting practice'. Critically, we need TI if we want to address wicked problems: we need diverse views in order to understand the problem (problem-setting), and to explore and develop solutions (solution-finding). Preferably, this is done in an iterative fashion (Dorst 2015; Steen 2013a). Simple as it may sound in theory—to put people from different disciplines in a project team—it can, however, be rather challenging in practice. Our aim is to better understand how we can organize TI effectively.

In the next 4-5 years, we will organize TI in three multi-partner innovation projects: ELSA Lab Defence; AI-MAPS ELSA Lab; and TRANSCEND. These projects deal with the design and deployment of AI systems in the domain of public safety and security; the people involved will draw from, e.g., Value Sensitive Design (Friedman and Hendry 2019) and Responsible Innovation (Von Schomberg and Hankins 2019) (ELSA stands for including Ethical, Legal, and Societal Aspects). In addition to these projects' particular research goals, we plan to conduct additional research, to address the question: What are effective ways to organize Transdisciplinary Innovation (TI)?

We plan to conduct longitudinal, participatory action research (McPhee, Hoppe, and Lindhult 2019); this will involve the organizing and evaluating of various interventions to promote TI. The differences between the projects offer opportunities to study different aspects of collaboration:
• in ELSA Lab Defence, we focus on collaboration between different types of organizations, e.g., government, academia, industry, and society—also, we plan to study the role of power;
• in AI-MAPS ELSA Lab, we focus on integrating expertise on Ethical, Legal and Social Aspects with expertise in tech—we plan to focus on method for practical and effective collaboration;
• in TRANSCEND, we focus on citizen engagement: to enable citizens to participate actively and creatively in the design and deployment of systems that they are supposed to benefit from.

We will use a systems view on TI, to look at the macro-level of a project's context, the meso-level of a project's content, and the micro-level of a project team's dynamics (Steen et al. 2021). We will also pay attention to the role of boundary objects (Star 2010): e.g., prototypes or experiments that people work on collaboratively. We may also draw from virtue ethics, to study the role of curiosity, creativity, and reflexivity (Steen 2013b; Steen, Sand, and Van de Poel 2021).

In this conference paper, we plan to discuss first findings from our observations from the first meetings of these three projects (October-December 2022); these first meetings are significant in that they help to establish the practical ways of working in these projects—notably with regards to processes and interventions to bring together and integrate expertise from different disciplines.

Session 11C: Knowledge for Practicing Engineers and Philosophers: Reconsidering Walter Vincenti's What Engineers Know and How They Know It

Zachary Pirtle, Claudia Eckert, Rémi Gandoin, Nina Jirouskova, Michael Poznic, Beth-Anne Schuelke-Leech, Martin Stacey & Loretta von der Tann

In 1990, the historian and aerospace engineer Walter Vincenti wrote What Engineers Know and How they Know It? which remains well-known in the history and philosophy of engineering. Vincenti argues that engineers develop unique categories of knowledge that are sometimes overlapping, and sometimes clearly distinct from the sciences. The book's legacy is prominent but mixed: while Vincenti's work is widely cited by policy and philosophy scholars as a reference on the nature of engineering (Houkes 2009, Kant and Kerr 2018, Zwart forthcoming, McKelvey and Saemundsson 2018, Nelson et al 2018), few engineers build upon his work. Arguably, there has not been a comparably deep book focused on engineering knowledge since Vincenti.

We will review Vincenti's categories of knowledge drawing upon both engineering and philosophical perspectives, touching on what seems plausible and useful from these views. We represent an interdisciplinary group, consisting of practicing engineers, engineering and design scholars, and researchers in philosophy of science and engineering. We will summarize Vincenti's key claims about categories of engineering knowledge . Importantly, we will highlight how Vincenti's research was strongly shaped by the work of Thomas Kuhn's and Rachel Laudan's work on scientific paradigms. We will also explore avenues to reassess the current validity of Vincenti's knowledge categories. We agree with some critiques of the conceptual haziness associated with some of Vincenti's knowledge categories, including his idea of design instrumentalities, and will discuss ways to improve them, touching on mathematical, design, and engineering knowledge. We also question to what extent computer modelling and simulation might create new forms of engineering knowledge, playing a crucial role beyond physical testing that Vincenti discusses.

For practicing engineers, something still resonates in what Vincenti is saying and the discussion on knowledge categorizations opens interesting avenues to explore improvements of current frameworks about how engineering works. It invokes reflections about the way "knowledge" relates to the current challenges of horizontal and vertical integration related to engineering (Barrella and Watson 2016); as well as about the value of engineering in multidisciplinary work environments.

We will conclude by considering ways in which deeper study of engineering knowledge might be important to engineering practitioners, including studying which types of knowledge might be more useful in developing new engineering systems (Pirtle and Moore 2019). We will include a special focus on how engineering methods do and do not fit into Vincenti's categories.

Session 11D: Responsibility in Practice

Zoë Robaey

The concept of responsibility is widely discussed in ethics of technology. From describing the different ways of conceptualizing responsibility, to determining conditions of responsibility ascription in different settings, to discussing distributions of responsibility and also responsibility gaps, the field abounds of helpful concepts. Empirical research has revealed multiple patterns of responsibility practice in research and technology development, ranging from producing sound science, over taking care of colleagues, to producing socially relevant technology. More recently, many ethicists of technology call for engaging with such practices for the sake of their improvement or normative steering. However, much of the engagement remains at the level of description and analysis of what is happening.

How can we move from critical analysis to transforming responsibility practices?

There are two key concepts that need attention: practice and transformation. In this session, we explore these concepts and their relation in more depth. Engaging with practice invites the ethicist, traditionally a conceptual thinker to engage with the empirical world where things can get muddy. Furthermore, claiming to transform said practices invites skepticism as it is unclear what the measure of transformation is and who is the judge of it.

In this session, we combine disciplinary perspectives to reflect on these concepts. Building a bottom-up approach for virtues for responsibility, Zoë Robaey presents the essential ingredients ethicists should consider when studying virtues in practice building on the work of MacIntyre and his critics. Presenting empirical methods in engaged ethnography, Mareike Smolka shows how these can enrich STS approaches such as the STIR method for transforming practices. Drawing from ethics and education and learning sciences, Thijs Loonstra sketches the main ingredients at the intersection of ethics and transformative learning that can provide a basis for further understanding what transformation requires. Reflecting on his own practice, Enrique Asìn Garcìa presents how through his work in experimental synthetic biology, he turned to empirical and conceptual investigations into value change. Finally, Philipp Neudert presents research on responsibility ascriptions and negotiations in a high-tech project, where the role of the engaged scholar is key.

Together, the presentations and discussions reflect on what it means to transform practices and whether this really leads to more responsibility in practice: engaged scholars, reflective practitioners, transformative education and expectations of responsibility are examined.

Session 11E: Value Change and Engineering

Ibo van de Poel & Freek van der Weij

1. Assignment
This workshop is intended to familiarize participants with different methods for dealing with value change in design and engineering. In the white paper of the ERC project 'Design for Changing Values', shared alongside this document, a variety of these methods are described. After the introductory talk by Ibo van de Poel, we ask you to form groups of 3-5 people. With your group, you will be asked to make a poster describing how one or more of the methods for dealing with value change should be applied to the development of **self-driving cars**.

Step 1: Select and develop a scenario where value change might occur (see section 4 for ideas)
Step 2: Propose how one or more of the methods mentioned in the white paper should be used to account for/analyze this type of value change. The goal here is to develop a proposal that could be executed in practice. It should not be overly ambitious, and choices for selecting a method should be justified. We have a panel of five experts who are here to answer your questions regarding a particular method for dealing with value change, see below.
Step 3: Provide proposals for design requirements based on your investigations.

2. Expert panel

Value monitoring	Tristan de Wildt
Technomoral scenarios	Olya Kudina
Prototyping	Joost Mollen
Simulation methods	Anna Melnyk
Design principles	Ibo van de Poel

3. Time schedule

11:00 - 11:20	Introduction
11:20 - 11:25	Group formation
11:25 - 12:10	Exercise
12:10 - 12:30	Feedback on posters

Session 12: From Causes to Actions

Samantha Kleinberg

The collection of massive observational datasets has led to unprecedented opportunities for causal inference, such as using electronic health records to identify risk factors for disease. However, our ability to understand these complex data sets has not grown the same pace as our ability to collect them. While causal inference has traditionally focused on pairwise relationships between variables, biological systems are highly complex and knowing when events may happen is often as important as knowing whether they will. In this talk I discuss new methods that allow causal relationships to be reliably inferred from complex observational data, motivated by analysis of intensive care unit and other medical data. Causes are useful because they allow us to take action, but do these models actually help us make everyday choices? In the second part of this talk I discuss just how people fare when using the output of machine learning and how we can go from data to knowledge to decisions, and what it means for inferring causal models that are useful and usable.

Session 13A: The Limits of Co-creation? The Ethics of Industry-funded Student Teams at Engineering Universities

Lukas Fuchs & Gunter Bombaerts

Engineering universities are increasingly interested in building direct collaborations with industry and other societal stakeholders. For example, student teams working on technical, entrepreneurial or societal problems are supported both by their university, as well as industry partners. Ideally, such student teams form an environment for students to work independently on relevant issues they are passionate for, apply their newly acquired technical and scientific knowledge and build a valuable professional network. Furthermore, such student teams may be a crucial way in which the boundary between universities and its ecosystem becomes more porous, allowing the flow of people and knowledge and adding to the effort to embed universities within society.

However, there are also ethical issues that have been raised in the context of some such student teams. Some student teams may not be funded by various companies, but may depend entirely on only one major company for funding. As a result, the company may have significant leverage in determining the set-up and goals of the student team. Financial and non-financial incentives may induce student teams to act as subsidiary organisations of the funding company. This may not be problematic if the student team is an entirely independent organisation. However, if the student team is associated with the university, then there is a substantive ethical question about where to draw a boundary between university and industry.

The article will make three distinct contributions. First, it will address questions about societal responsibility of universities (Collini 2012) and the role of engineering universities in particular (Miller 2019). Second, by discussing issues in co-creation with industry, it adds to the debate about worries of academic capitalism, academic entrepreneurship (Siegel & Wright 2014) and co-creation for sustainability (Treicher et al. 2014). These debates have usually focused on large-scale societal trends in higher education, but we contend that important insights are gained by considering small-scale ethical questions that deal with the terms and modes of co-creation. Finally, by focusing on student teams, we can add to the recently emerging debate on the values in university-industry collaboration (Hillerbrand & Werker 2019).

The central part of our discussion will revolve around discussing the case of a student team at Eindhoven University of Technology (TU/e). This student team was solely funded by one international company which determined the set-up of the student team, its activities and goals, as well as its leadership. The company is a powerful international corporation and there are serious concerns about the societal legitimacy of its activities. The ethical review board of the university was tasked with deciding whether this student team should be officially recognised. After a positive decision, the team receives visibility on the university website, has access to certain resources and training and is allowed to use the university logo. However, in summer 2022, the ethical review board advised against recognising this particular student team. In the paper, we describe this case and aim to analyse the ethical considerations behind the decision of the ethical review board.

Session 13A: Value Change by Value Elicitation, Value Amplification, and Value Integration in Interdisciplinary Ethics Education: Examples from the Berlin Ethics Certificate

Sabine Ammon, Alexandra Kljagin, Juliane Rettschlag & Martina Vortel

To catalyze value change of existing sociotechnical systems is a complex process. Our contribution looks at one puzzle piece in this endeavor, namely, the instigation of personal values of future innovators which comes with a changed attitude towards values in the professional context. It discusses an educational framework as well as examples of ethical interventions of a certification program, the Berlin Ethics Certificate, which aims to ethically empower students from a technical university.

Although recent research in Value Sensitive Design, Design for Values and Responsible Innovation has shown that the thesis of the value-neutrality of technology is wrong and cannot be upheld, it is still persistent. This assumed value-neutrality of technology has a corresponding assumption, namely, that professional work should refrain from incorporating personal attitudes and values. Students as future innovators should contribute only their disciplinary knowledge to the design and development process but are not to be present in this process as a person with an ethical value compass. A change of this technocratic understanding of technology is not only a change of the design and development process of artifacts, but also a change of how designers, developers, and innovators contribute to the emergence of technology.

The Berlin Ethics Certificate program aims to change this. Future innovators can best contribute, so our assumption, to value change if they are aware of their own ethical values and their personal attitudes as a situated person also in their professional work and if they are aware of how these contribute to collectively shared values. To achieve this, the Berlin Ethics Certificate works with the concepts of value elicitation, value amplification, and value integration. By value elicitation we understand the process of making the students aware of their attitudes, their ethical values and their ethical agenda, e. g. by discussing real cases (rather than abstract ethical principles) so that students are able to reason based on their personal stakes in the matter. Jointly exploring and grounding their attitudes, values, and visions in discussion, critical course work, and projects support them in a process of value amplification in which their ethical positioning is maturing. By value integration we encourage the students to integrate their ethical agenda into their professional work. In our contribution we will discuss this conceptual framework and will show examples from teaching practice.

Session 13B: Toward engineering codes of conduct for philosophy of technology

Pieter Vermaas

In this contribution I explore the basis for and content of a code of conduct for philosophy of technology. The need to have codes of conduct in engineering and the analysis of their content, have been topics of research in ethics of engineering, positioning philosophy of technology outside of engineering, in line with the more traditional reflective and critical stances philosophy of technology adopted toward technology. Yet, with the emergence in the last few decades of a more constructive relation between ethics of technology and engineering as witnessed by design for values and ethics by design (e.g., Van den Hoven et al. 2015; Dainow and Brey 2021), it can be argued that philosophy of technology has become in part engineering, leading to the question of whether it should also accept these studied codes of conduct for itself.

When approached logically the answer to this question should be unconditionally positive. Yet, in practice acceptance of this conclusion may be harder. In design research, under headings as value-sensitive design and social design (e.g., Friedman and Hendry 2019; Tromp and Hekkert 2019), a similar development has taken place to employ design for the good. And in such responsible design it seems that codes of conduct accepted for engineering design and design research can easily be suspended (Vermaas 2019).

In this contribution I adopt the logical position and move to the subsequent question of content a code of conduct for philosophy of technology could have. First I explore what parts of engineering codes of conduct can apply to philosophy of technology that constructively contributes to the design and development of technology. And, using the above degression to design research, I extend this exploration to design research codes of conduct. Second I open the floor to discuss with you the need for a code of conduct for philosophy of technology.

References:

Dainow, B., and P. Brey (2021) Ethics by Design and Ethics of Use Approaches for Artificial Intelligence, Version 1.0, European Committee, https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/ethics-by-design-and-ethics-of-use-approaches-for-artificial-intelligence_he_en.pdf
Friedman, B., and D.G. Hendry (2019) Value Sensitive Design. MIT Press.
Tromp, N., and P. Hekkert (2019) Designing for Society: Products and Services for a Better World. Bloomsbury Visual Arts.
Van den Hoven, J., P.E. Vermaas and I. van de Poel (eds.) (2015) Handbook of Ethics, Values and Technological Design, Springer.
Vermaas, P. (2019) Transparency in Responsible Design: Avoiding Engineering Overconfidence and Supporting Societal Acceptance, in Proceedings of the Design Society: International Conference on Engineering Design, 1(1), 3431-3440.

Session 13B: Engineering codes of ethics with care

Diane P. Michelfelder & Sharon A. Jones

How might the ethical codes adopted by professional engineering societies in the United States benefit from being integrated with and grounded on an ethic of care? In forming and taking steps toward answering this question, in particular by appealing to work in feminist care ethics by Carol Gilligan (1982), Nel Noddings (2002, 1984), Joan Tronto (1993) and Daniel Engster (2007), Elisa Warford (2018) has suggested that to incorporate care into engineering codes of ethics would also be to put an emphasis within the codes of the importance of being aware of the context within which engineering design decisions need to be made so that negative outcomes such as increasing the vulnerability of already-vulnerable peoples can be avoided.

Since the publication of Warford's paper, the upsurge of interest among engineers in re-envisioning engineering activities from engineering education through engineering design by looking at them through the lens of the ethics of care has continued to trend upwards (see for instance Baas et al. 2022; Fore et al. 2020; Nair and Bulleit 2019; Russell and Vinsel 2019). In the past five years, several US engineering societies have also revised some of the canons of their codes of ethics (ASME 2021; ASCE 2020; IEEE 2020) in ways that raise the commitments of these codes to sustainable development, though not necessarily to care. These two trajectories have proceeded along parallel tracks. The intent of this paper is to bring them together.

Specifically, we aim to bring them together by expanding on Warford (2018) to imagine how grounding engineering codes of ethics within a context of care ethics might elevate the commitment of engineers to sustainable development, both with respect to codes that have already made positive shifts in this direction and to ones that haven't yet (e.g., AIChE, NSPE). We first suggest why it makes sense to do this given the realities of the significant issues facing society in the 21st century. From there we go on to suggest potential changes to the language of the codes that could make them better grounded in the ethics of care. The paper ends by acknowledging and responding to objections that could be made to these transformational changes.

References:

Baas, L., Metselaar, S. and Klaassen, P. (2022). Circles of Care for Safety: A Care Ethics Approach to Safe-by-Design. Nanoethics 16: 167–179. https://doi.org/10.1007/s11569-022-00419-w.
Engster, D. (2007). The Heart of Justice: Care Ethics and Political Theory. Oxford University Press.
Fore, G.A, Hess, J. L, and A. Katz. (2020, June). Ethics in Engineering or Engineering in Ethics? Paper presented at the ASEE Annual Conference and Exhibition, Montreal, Canada.
Gilligan, C. (1982). In a Different Voice: Psychological Theory and Women's Development. Harvard University Press.
Nair, I. and Bulleit, W.M. (2020). Pragmatism and Care in Engineering Ethics. Science and Engineering Ethics 26 : 65–87. https://doi.org/10.1007/s11948-018-0080-y
Noddings, N. (2002). Starting at Home: Caring and Social Policy. University of California Press.
Noddings, N. (1984). Caring: A Feminine Approach to Ethics and Moral Education. University of California Press.
Russell A. and Vinsel L. (2019). Make Maintainers and an Ethic of Care. In M. Wisnioski, E. H. Hintz, and M. S. Kleine (eds.), Does America Need More Innovators? MIT Press, pp. 249-269.
Tronto, J. (1993). Moral Boundaries: A Political Argument for an Ethic of Care. Routledge.
Warford, E. (2018, June). Toward a More Caring Code of Engineering Ethics. Paper presented at 2018 ASEE Annual Conference and Exposition, Salt Lake City, Utah.

Session 13C: The Environmental "Paradox of Prevention" in Socially Disruptive Medical Technologies

Cristina Richie

The carbon emissions of global healthcare activities make up 4-5% of total world emissions, placing the healthcare industry on par with the food sector. Yet, the environmental impact of health care has been underconsidered, in part, because of the belief that all available health care technologies are medically necessary and therefore carbon emissions are morally irrelevant.

Many of emerging and socially disruptive technologies relate to clinical services that support and protect longevity, health, and disease prevention and treatment, such as the digital twin and personalized medicine. Hospital care and physician and clinical services are the two largest carbon contributors to health care—exceeding even healthcare structures. While significant work is already being done on carbon reduction in health care and sustainable technologies, less attention has been paid to sustainable health care technologies and even less to the "paradox of prevention"—the idea that preventive health care technologies may extend lifespans and thus increase the carbon of health care both in an individual's life and in the medical industry overall.

This presentation will, first, overview the importance of sustainability in emerging and socially disruptive technologies, focusing on the intersection of technological and biomedical ethics. It will, second, examine current socially disruptive technologies in preventative health care and third, present the "paradox of prevention." The conclusion will contemplate how ethicist might think about sustainability in preventive health care technologies in the future.

Session 13C: Turn, Turn, Turn: Wind Turbines, Aesthetic Persuasion, and Changing Feelings

James Hutton

Onshore wind turbines present an opportunity for cheap, abundant clean energy. Yet in some countries, the build-out of onshore wind has stalled due to aesthetic objections from citizens. What is to be done? I argue that a strategy of aesthetic persuasion, in which we work to promote positive aesthetic feelings about wind turbines, is both feasible and ethical.

To motivate aesthetic persuasion, I first consider two rival strategies:

Strategy 1: Acquiesce. The first option is for governments and firms to treat citizens' aesthetic objections as authoritative and to ban onshore wind turbines (Scruton 2012). Clearly, this should be a last resort, because the opportunity costs of abandoning onshore wind are extremely high, both in terms of economics and human wellbeing.

Strategy 2: Overrule. The second option is for governments and firms to overrule citizens' aesthetic objections and to build onshore wind turbines anyway. This has some merits, but seems somewhat undesirable too. It's undemocratic to overrule citizens' aesthetic feelings without engaging with them. Moreover, this approach threatens to breed resentment, risking a future backlash.

Strategies 1 and 2 treat people's aesthetic feelings as fixed. However, as art critics and advertisers know, it's possible to shift people's sense of what's ugly or beautiful through aesthetic persuasion. If I want to convince you that an object is beautiful, I can guide your attention to features you haven't noticed; I can encourage you to draw associations between this object and others you admire; and I can supply relevant information about its merits (e.g., its functionality). I illustrate these modes of aesthetic persuasion by analysing recent efforts by the gas industry to shaped consumers' aesthetic feelings about gas stoves (Leber 2020). Next, I argue that it's feasible to reshape citizens' aesthetic feelings about wind turbines by highlighting the sleekness of wind turbines, their dynamic power, the biomimicry of their blades, the positive vision for the future they embody, etc. Plenty of citizens already hold that wind turbines are not ugly (Klæboe and Sundfør 2016) and engineers have long taken aesthetic considerations into account when designing turbines (Owens 2019), so it's feasible to persuade citizens that wind turbines are not ugly.

Aesthetic persuasion is feasible, but is it ethical? We might worry that it's manipulative to try to alter citizens' feelings. However, drawing on the idea of responsiveness to aesthetic reasons (Cross 2017), I argue that aesthetic persuasion is not manipulative so long as it involves an honest attempt to highlight reasons for finding wind turbines beautiful.

I close by identifying further investigation that's needed to flesh out the aesthetic persuasion strategy and the broader applications it might have.

References

Cross 2017. 'Art Criticism as Practical Reasoning'. British Journal of Aesthetics 57 (3): 299–317.
Klæboe and Sundfør 2016. 'Windmill Noise Annoyance'. International Journal of Environmental Research and Public Health 13 (8): 746.
Leber 2020. 'The Gas Industry Is Paying Instagram Influencers to Gush over Gas Stoves.' Mother Jones. 2020.
Owens 2019. The Wind Power Story. Hoboken: Wiley.
Scruton 2012. Green Philosophy. London: Atlantic.

Session 13D: Designing with nature: reflections on biomimetic practices

Alessio Gerola, Zoe Robaey, Vincent Blok, Dayo Jansen, Jaco Appelman & Paul Breedveld

Biomimetic design (from the Greek bios, life, and mimesis, imitation) consists in the conscious emulation and integration of biological models in technological design with the aim of solving today's technical and ecological challenges. For example, the water-repellent properties of lotus leaves have inspired self-cleaning materials and paints. Numerous different approaches to biomimetic design currently exist, fragmenting the field in a myriad of approaches characterized by partial overlaps and occasional contrasts. The rich but often inconsistent nomenclature reflects this patched framework: biomimetics, biomimicry, bionics, bio-inspired design etc. represent some of the most common terms proposed to organize the field. Making sense of this fragmentation, however, demands further reflection on the scope, goals, and methods of biomimetic design applied to different domains, such as bio-inspired robotics and engineering, biomimicry approaches for sustainable architecture etc. Different approaches are in fact characterized by different assumptions about the conceptual and normative meaning of nature and of its emulation. In this panel we present a philosophical approach to understanding different types of biomimetic design and their underlying assumption, inviting engineers and designers to present their work and discuss the current status of the field of biomimetic design. Together we will discuss and reflect on the following questions:

- What does it mean to mimic a natural model? What methods are employed?
- Why are biomimetic designers trying to imitate nature?
- What aspects of nature, of an organism or ecosystem, should be imitated, and why?
- What does it mean to design with nature in different biomimetic domains?
- What are the promises pushing for the development of biomimetic design?

Panel organizer: **Alessio Gerola** (Wageningen University & Research)

What philosophy can bring to engineers & designers

Alessio Gerola, Zoe Robaey, Vincent Blok: *What does it mean to mimic nature? A typology for biomimetic design*
Wageningen University & Research

In an effort to produce new and more sustainable technologies, designers have turned to nature in search for inspiration and innovation. Biomimetic design is the conscious emulation of biological models to solve the technical and ecological challenges of today. Nowadays numerous different approaches exist that study and mimic nature, a variety of practices that comes in turn with different and often unacknowledged assumptions about the meaning and value of nature and of its emulation. How is it possible to make sense of the fragmentation of the broad field of biomimetic design by looking at the conceptual and normative assumptions that divide its different approaches? In order to make sense of the current fragmentation of the field, the paper provides a conceptual typology that helps characterizing different biomimetic approaches. It does so by articulating and analyzing existing assumptions about the conceptual and normative dimensions of mimicry of nature. Their intersection generates six different biomimetic types, whose relations and assumptions are discussed. The typology illustrates the conceptual and normative tensions that are driving different research strands in biomimetic design.

Alessio Gerola is PhD candidate in Philosophy and Ethics of Technology at the Philosophy Group of Wageningen University, supervised by dr. Zoe Robaey and prof. Vincent Blok. His research explores the conceptual and normative implications of biomimetic design on the meaning of nature, sustainability, and human relations to nature.

What engineers & designers can bring to philosophy

Dayo Jansen: *Understanding animal flight and escape to create and improve bio-inspired robots*

Wageningen University & Research

Animals have evolved novel techniques for interacting with their surroundings during flight. These techniques have inspired engineers and designers to develop human-made devices, structures, and technologies. However, many times the translation of biological solutions to bio-inspired designs lacks input and knowledge from biologists. In our group's research, we aim to fill this gap by investigating the functional morphology and locomotion of animals using a quantitative approach to reveal underlying physical principles that can be translated to bio-inspired designs. Dayo's PhD focuses primarily on the interaction between bats and moths. By doing this he aims to understand what makes bats successful

in catching moths and what moths do to avoid getting captured at all costs. He translates these principles to technical designs like autonomous drones that hunt down moths in greenhouses during the night, just like bats would, and reduces the need for conventional pesticides to deal with these pests in greenhouses. Our approach is mostly experimental with mathematical modelling and numerical simulations to support the experiments and phylogenetic analyses to interpret our findings within biological and ecological contexts. The research questions we aim to address are: What role do specialized flying manoeuvres exerted by bats and moths play in successful hunting attempts? How can we translate what we learn from biology to designs and technologies?

Dayo Jansen is a second year PhD student at the Experimental Zoology group at Wageningen University and Research and collaborates with PATS indoor drones solutions at the faculty of Aerospace Engineering . Through his research, he hopes to address many questions concerning the interaction between predator hunting and prey escape , as well as motivate bio-inspired solutions for interfacing within complex environments to reduce the need for conventional harmful pesticides.

Jaco Appelman: *Bio-inspired Innovation*

Utrecht University

In 2020 it was the first time that humans produced more in terms of products and waste also called technomass than Earth produced biomass. Our society strongly relies on science and technology to solve our problems and to fulfil our needs. We are becoming increasingly aware of the fact that science and technology are not enough to provide us with a society that nourishes our surroundings, but rather the opposite. Our products and innovations consume more natural resources that they create. We have designed both our world economy and product based on the idea that nature provides us with costless materials in abundance.

Over time, we have started to treat nature as a dump that can store and clean up all our newly created waste. To become a society that takes care of our Earth, and thus ourselves, we need to switch to circular systems.

Instead of designing these new systems ourselves, we could look at existing systems that are circular and harmless by nature by taking inspiration from nature. What can we learn from nature, what lessons do we internalize as a species? Can you isolate aspects and still create a system that adheres to the operating conditions of Earth? Bio-inspiration is a relatively new field within innovation, which is all about creating innovative ideas by looking at and learning from nature.

Jaco Appelman is co-founder and programme coordinator of the Bio Inspired Innovation programme at Utrecht University, where he teaches and conducts research into bio-inspired change.

Paul Breedveld: *BITE – Bio-Inspired Technology Group*

TU Delft

Collaborating with biologists, medical companies and (academic) medical centres, the research within Breedveld's research group BITE (Bio-Inspired Technology) has resulted in a great number of innovative medical devices, such as multi-steerable instruments and catheters inspired by anatomy of squid tentacles, adhesive instruments and self-propelled intestine devices based on starfish and tree frog adhesion, highprecision biopsy harvesters inspired by chewing organs of sea-urchins, mechanical follow-the-leader instruments inspired by snakes and mechanical calculators, self-propelled steerable needles and tissue transporters based on ovipositors of parasitic wasps, and integrated-assembly 3D-printed instruments and prostheses designed for low-cost use in developing countries.

Breedveld's research was rewarded in 2012 with a prestigious Dutch VICI research grant on the development of advanced snake-like instrumentation for endo-nasal skull base surgery and in 2013 with an Antoni van Leeuwenhoek distinguished professorship at TU Delft. In 2019 Breedveld became one of the chairs of Dutch Soft Robotics: a large national research programme of the Dutch 4TU Federation aimed at using clever bioinspired solutions for the development of medical and agricultural soft robotic systems.

Session 13E: AI-enabled decision-support systems: A catalyst or curse for human's ethical decision-making?

Franziska Poszler & Benjamin Lange

Emerging technologies entail comprehensive transformations across society that even touch on individuals' reasoning, decision-making processes and morality (e.g., Chan et al., 2020). Corresponding technological applications are, for example, 'Robo-advisors', virtual assistants or clinical decision support systems for healthcare professionals (Erler & Müller, 2021). Similarly, scholarly debates and investigations of the technological mediation of morality are picking up (e.g., Verbeek, 2006), but are not yet conclusive about the direction of influence. On the one hand, corresponding technology may allow individuals to make more enlightened, value-aligned and rational decisions; on the other hand, increased reliance on technologies for ethical decisions may leave individuals morally deskilled in the long-run (Vallor, 2015). This highlights the prevailing ambiguity and range of influences of technology on human ethical decision-making.

To understand the full spectrum of technologies' influences on humans' ethical decision-making, this article provides a structured review and analysis of 33 pertinent publications. Main takeaways are identified underlying sources (i.e., agent and system), mechanisms and corresponding design features (i.e., process-oriented or outcome-oriented navigation) as well as outcomes (i.e., deliberation, motivation, autonomy and action enhancement) that arise from human interaction with AI-enabled ethical decision-support systems. An integrated framework is derived that establishes links between the identified sources, mechanisms and outcomes of influences and that proposes seven interrelations. For example, it is postulated that a technology's contribution to a users' deliberation enhancement decreases when moving from process-oriented to outcome-oriented navigation. Apart from these influences on an individual level, long-term implications for society (i.e., changing professions/duties, humanity's infantilization, liability issues and restricted moral progress) are highlighted.

Overall, this article aims to generate important theoretical and practical insights for the broader public, practitioners and scholars. Namely, the broader public will be informed about potential (subconscious) influences of decision-support technologies on their ethical decision-making. For practitioners such as technology companies, opportunities and risks of influences are sketched that may need to be contemplated/addressed during the development process. For the scientific community, the derived integrated framework and proposed interrelations can serve as a baseline for future empirical studies investigating and validating potential sources, drivers and outcomes that relate to technologies' influence on ethical decision-making.

References

Chan, L., Doyle, K., McElfresh, D., Conitzer, V., Dickerson, J. P., Schaich Borg, J., & Sinnott-Armstrong, W. (2020). Artificial Artificial Intelligence: Measuring Influence of AI 'Assessments' on Moral Decision-Making. Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society (pp. 214-220).
Erler, A., & Müller, V. C. (2021). AI as IA: Human enhancement through artificial intelligence (AI) for intelligence augmentation (IA)?. In M. Iena & F. Jotterand (Eds.), The Routledge handbook of the ethics of human enhancement. Routledge Taylor & Francis Group.
Vallor, S. (2015). Moral deskilling and upskilling in a new machine age: Reflections on the ambiguous future of character. Philosophy & Technology, 28(1), 107-124.
Verbeek, P. P. (2006). Materializing morality: Design ethics and technological mediation. Science, Technology, & Human Values, 31(3), 361-380.

Session 13E: AI-enabled decision-support systems: A catalyst or curse for human's ethical decision-making?

Franziska Poszler & Benjamin Lange

Prompting meaningfulness via value-centered technological interactions is critical if we wish to create technology that supports human flourishing (Bynum, 2006). Currently, privacy decision-making - whether it be on our phones or online—is largely lacking in meaning. When we "click through" an overwhelming amount of privacy notifications or are nudged by dark patterns to consent to giving away our data (Brignull, n.d.), we are acting impulsively, without due reflection on who we are and what we value.

Here, we will discuss recent findings from a qualitative, semi-structured interview investigation exploring how we value data privacy. This study is part of an interdisciplinary project investigating the role of personal values in data privacy decisions as a means to promote more meaningful user privacy decisions.

In the interviews, questions probed the relationship between privacy and values when selecting a smartphone application as well as privacy decisions made more generally (e.g., online). To elucidate the often unconscious values involved in privacy decisions, questions were designed to obtain statements pertaining to commitments that represent values (example: “can you walk me through how you go about choosing a smartphone app to download?”). This is primarily drawn from an understanding of values as clusters of commitments pertaining to how to be or act that share a similar end state or object (Killmister, 2017) [1]. In addition to probing commitments, participants were also asked about their affective responses as a representation of underlying values (example: “Can you think of a time when something felt “off” about your privacy online or on your phone?”). Coding proceeded using a process of reflexive thematic analysis (Braun and Clarke, 2022), with emphasis on instances where participants felt unable to act according to their values or where values appeared to be in tension.

We will compare the results of this investigation with an earlier quantitative investigation that surveyed users and their privacy preferences (manuscript in progress). Limitations resulting from the mostly WEIRD (Western, Educated, Industrial, Rich, Democratic) participants will also be discussed (Henrich et al., 2010). We conclude by contextualizing this investigation into individual values and privacy decisions within broader societal norms and values around data privacy.

[1] For example, consider an agent who values personal security. Her value of personal security involves commitments not to share pictures of herself online and to not share her location with smartphone apps, both oriented towards the desirable end state of maintaining her personal security.

References

Brignull, H. (n.d.). Deceptive design. Retrieved September 29, 2022, from https://www.deceptive.design
Bynum TW (2006) Flourishing ethics. Ethics Inf Technol 8:157–173. https://doi.org/10.1007/s10676-006-9107-1
Braun V, Clarke V (2022). Thematic analysis: a practical guide. (1st ed.). Sage.
Henrich J, Heine SJ, Norenzayan A (2010) The weirdest people in the world? Behav Brain Sci 33:61–83. https://doi.org/10.1017/S0140525X0999152X
Killmister, S. (2017). Taking the measure of autonomy: A four-dimensional theory of self-governance (1st ed.). Routledge.

Session 13E: The Interconnection of Subjective Values and Experience-Based Learning

Oyku Ulusoy

How do subjective values change, when an individual learns something fundamentally new? I investigate the experience of value change as an outcome of new discoveries by drawing on Transformative Experience (Paul, 2014). By subjective evaluation, I describe the comparative evaluations individuals make from their own point of view to rank their preferences over alternatives they face in their lives. I argue that new technological implementations are transformative phenomena that lead to significant alterations in individuals’ understanding, which lead to creating of new knowledge and new subjective values. I also argue that one's current understanding, which is constructed through past knowledge and existing values, functions like a scientific paradigm (Kuhn, 1996) and shapes individuals’ response to new discoveries, how they experience value changes and how their preferences are altered as they adjust to the changes.

Paul (2014; 2015) provides a compelling philosophical framework to depict how learning something fundamentally new alters someone’s perspective, i.e., their knowledge and understanding, and as a result lead to changes in their comparative evaluations. For this argument, she draws on Frank Jackson’s (1982) thought experiment, Mary from the Black and White Room. Mary is a neuroscientist who lives in a black and white room, where she has never seen colours. Paul depicts that seeing red is a transformative experience, which provides new phenomenal content to Mary, which then alters her understanding. After Mary sees red, she learns how to imagine situations that involve red. She can then know how she values these situations. Paul argues that it is necessary for individuals to be able to grasp and understand new phenomena from their own perspective in this way to be able to know what they value and prefer. Correspondingly, Paul argues that transformative experiences reveal a sort of epistemic limitation in understanding. Individuals cannot be sufficiently advised about new phenomena before experiencing it themselves to make their own and informed evaluations about the alternative outcomes of the phenomena.

I argue that policy-shifts in response to unknowns such as Covid-19 and the climate change and new technological implementations result in a similar situation of epistemic unpreparedness for individuals. New policies and changes of technology can be transformative experiences leading to value changes, where understanding is limited, and individuals are unprepared to know how desirable they find the novel scenarios with the new phenomena. Individuals then are limited to grasp the extend of the change to respond by adjusting their understanding, values and choice-relevant preferences.

I also argue that exposure to new phenomena is not sufficient to experience changes in understanding and values as Paul depicts. Individuals ought to have prior knowledge to be able to grasp the new content and adjust their perspective accordingly. History of scientific discoveries reveal cases, where scientists fail to realize they discove red something new, which ought to alter their understanding (Kuhn, 1996). Discoveries of X-Rays and Oxygen emphasize that there can be a level of prior understanding that can prepare individuals to accurately shape their expectations about new phenomena.

References

Jackson, F. (1982). Epiphenomenal qualia. Philosophical Quarterly, 32: 127-36.
Kuhn, T. (1996). The structure of scientific revolutions (3rd ed., Issr collection). Chicago, IL: University of Chicago Press.
Paul, L. A. (2014a). Transformative experience. Oxford University Press.
Paul, L. A. (2015). What you can’t expect when you’re expecting. Res Philosophica, 92(2), 149–170. https://doi.org/10.11612/resphil.2015.92.2.1

Posters

The ChatGPT hype as a medium to popularize and jeopardize social debates about values and virtues in education

Maximilian Rossmann

It is obscure which values matter in the assessment of unproven technology. After two months of intense debate about ChatGPT3 in the classroom, it seems unlikely that the structure and organization of education will disruptively change. The narratives and arguments were not entirely new, but OpenAI and ChatGPT3 provided props and thrilling playgrounds to make-believe different worlds and share fascinating stories, even without believing them themselves. This presentation examines the functions and linguistic features of “futuristic communication” (Grunwald 2013) and make-believe (Roßmann 2021) to discuss the motivations, emergence and consequences of the ChatGPT3 hype in education, and asks what a more responsible communication about technology futures and ethics might look like.

Hype and alarmism create hyperbolic levels of attention and expectations around issues and easily jeopardize health decisions, resources, and trust in science. ‘Hyping’ violates the norms of science communication by overstating the certainty, accuracy, and relevance of facts (Intemann 2020). Hyping, so to say, calls for the re-assessment of beliefs by unproportionally overstating the importance of uncertain evidence. Revealing and distinguishing different functions of future-oriented communication in technology development, however, allow us to explain what motivates both unintentional hype and strategic pretense: Most prominently, economic sociologies suggested that many future-oriented practices do not aim to factually realize a project but instead manipulate the market value of research assets, such as laboratory equipment, research networks, or topical expertise (Birch 2017). The Sociology of Expectation highlighted the coordination function, as stakeholders mutually observe and adjust their projected actions (Van Lente & Rip 1998). Hermeneutic Technology Assessment (Grunwald 2020) understands visionary communication as a relevant medium for social debates. And finally, the philosophy of imagination (Kind 2016) discussed the conditions and constraints for imagination to motivate action and impact beliefs at the danger of "imaginative illusions".

This presentation explains how hype can jeopardize ethical debates and technology assessment (see Nordmann 2007, Vinsel 2021). A focus lies on revealing how key narratives ink technology with values and virtues by representing features and uses of technology as deviations from normalcy demanding a response that is either punished or rewarded as the story progresses (Roßmann 2021). The presentation illustrates how hype dynamics work and accelerate by attention patterns and linguistic features of an inverted ‘Chinese whisper’ effect that helps to spread the most extraordinary stories which then call for a sobering critique that maintains their topicality, irrespective of demystifying arguments or evidence.

References

Intemann, K. (2020). Understanding the Problem of “Hype”
Grunwald, A. (2013). Techno-visionary sciences: Challenges to policy advice.
Van Lente, H., & Rip, A. (1998). Expectations in Technological Developments: An Example of Prospective Structures to be Filled in by Agency.
Grunwald, A. (2020). The objects of technology assessment. Hermeneutic extension of consequentialist reasoning.
Kind, A. (2016). Imagining Under Constraints.
Nordmann, A. (2007). If and Then: A Critique of Speculative NanoEthics.
Lee Vinsel (2021) You’re Doing It Wrong: Notes on Criticism and Technology Hype
Roßmann, M. (2021). Vision as make-believe: How narratives and models represent sociotechnical futures.

Roboticizing Emotions? A Framework for Analyzing Care Workers’ Emotions towards Assistive Robots

Belen Liedo

While emotions have not always been addressed in philosophy of technology, recent trends argue in favor of including them in the philosophical and ethical scrutiny of disruptive technologies. As Roeser convincingly states, emotions may tell us something important about our relationship with technology and its ethical significance (Roeser, 2018). When we talk about assistive robotics, we should keep in mind that care tasks are a complex phenomenon that includes specific ethical values and practical abilities (know-how), and that they are often developed within sets of relationships that forge through time. The perceptions, desires, personal values, and acceptance of the people involved need to be included in any ethical analysis of assistive robotics (van Wynsberghe, 2015).

Among other elements, emotions arguably play a key role in this ethical evaluation of assistive robotics in care environments. A care environment is an institutional or domestic locus in which care is performed, and it includes a set of relations between people and artifacts. When these artifacts are social robots, their impact on human relationships may challenge the success of good care. Here good care is defined, inspired by the classic work of Tronto, as the meeting of the relative wellbeing of all people involved in the five aspects of care (caring about, caring for, care-giving, care-receiving, and caring-with) (Tronto, 2013). From the framework of ethics of care, emotions are a key factor in the evaluation of this success.

In this talk, I shall focus on one specific objective. I propose to concentrate on the emotional responses of professional care workers when a robot is introduced in their care setting. I will argue that the emotions of care workers may help to evaluate the ethical significance of the robot’s presence in care relationships. More concretely, I will propose a framework for assessing the influence of the robot on the already existing human relationships. For doing so, I will argue that the presence of the robot might impact at least two main aspects of care work: communication and agency.

Regarding communication, robots may influence epistemic and affective trust, fluency, and the attunement of communication between care-receivers and caregivers. Regarding agency, robots may affect the understanding of situations, the capacity to taking free decisions, and the self-trust of care workers. I defend that these two aspects are determinant for the possibility of good care, and that good care is compatible with the introduction of an assistive robot as long as its impact on them is satisfactory.

References

Roeser, Sabine (2018). Risk, Technology, and Moral Emotions. Routledge.
Tronto, Joan C. (2013). Caring Democracy. Markets, Equality and Justice. New York University Press.
Van Wynsberghe, Aimée (2015). Healthcare Robots: Ethics, Design and Implementation. Routledge.

A “Data Trustee” as Mediator between the Value of Privacy and Data Analysis

Antonia Kempkens

Since Artificial Intelligence (AI) algorithms allow for data analysis on a grand scale, the value of privacy has changed. It is now in conflict with data analysis that uses personal information. I propose the introduction of a “data trustee” as a solution for the conflict between data analysis and privacy, because it enables the integration and modification of the value of privacy.

Privacy refers to the control over the access to and use of bodies, places and personal information, as in the ability to selectively open and close oneself to others (Moore 2010:27; Acquisti et al. 2022:270). Previously, the control over places and bodies was the overriding objective as all personal information could be found there. Nowadays, personal information is accessible from anywhere. The value of privacy has to adapt to technological developments by shifting its focus to the protection of personal information.

Using the example of health data, patients’ data could be analysed in order to support effective treatment of new patients. But patients’ data are private information. In order to respect the patients’ privacy, first, patients would have to be asked for their informed consent to the usage of their data and, second, the data would have to be anonymised before processing. By informing the patients, about what is to happen to their data and giving them the opportunity to consciously consent or decline, the control over access to their personal data is vested in the individuals. Moreover, if the data is not anonymised, the patient loses control of their personal information contained in the results of data analysis. That is why, without the patients’ informed consent and without anonymisation, data analysis violates the patients’ privacy.

For technology to be helpful, its design should integrate the value of privacy. I discuss the idea of a “data trustee” as a mediator between patients and doctors or researchers as a technical solution for the implementation of the privacy protection. Thereby, every patient concludes a contract with the “data trustee” which specifies which data may be used for which purposes. The “data trustee” gets all the data and manages who can access what data. The data remains with the "data trustee" and AI algorithms are allowed to learn from the anonymised data via federated learning. This prevents the storage and dissemination of personal information without the patient's control. Thus, the value of privacy and data analysis can influence and respond to each other as they change. Finally, I argue that it is worth taking time for the implementation and usage of the “data trustee”, because it protects the value of privacy and enables the patients’ trust in data analysis.

References

Acquisti, Alessandro; Laura Brandimarte und Jeff Hancock (2022): „How privacy’s past may shape its future - An account of privacy’s evolutionary roots may hold lessons for policies in the digital age“, Science, Vol. 375(6578).
Moore, Adam D. (2010): „Privacy Rights: Moral and Legal Foundations”, Pennsylvania State University Press.

Virtue Ethics and software development: debates and proposals

Olaya Fernández Guerrero, Ángel Luis Rubio García & Julio Rubio García

Ethical issues have a growing interest for software developers, in as much as technological developments are more and more present in our lives, and it has become evident that technologies are morally charged (Verbeek, 2014). From an interdisciplinary approach, gathering researchers from the fields of computer science and moral philosophy, this paper reflects on virtue ethics –first defined by ancient Greek philosopher Aristotle- and explores the possibilities in which this classical moral notion can be addressed in contemporary contexts of software development (Haggendorf, 2020; Gamez et al., 2020). Together with general values such as interaction, correctness and security, widely discussed in the fields of Computer Science and Artificial Intelligence, the research also highlights the importance of intentionality. Given the fact that most software developers usually work in a team, and they often take items of software from different sources available, the implementation of any new software shall be understood as a complex process fragmented among many agents (with their consequent and perhaps hidden intentions), hence it is difficult to identify the bearer of moral responsibility if that software leads to unethical consequences. The paper elaborates on the notion of ‘intentional dependency network’ to name this chain of complex interactions in the context of software development. The nodes of these intentional dependency networks are constituted by three elements: a human being who interacts with a software artefact with a declared intention. Inconsistencies may arise in a node at local level, since a software artefact can produce certain effects that are incompatible with the intention declared by the human. Thus deepening in the complex ecology of current software development and its networks provides an interesting field to explore the origins and evolution of moral agency. Once thoroughly examined the topics above mentioned, the paper ends with a preliminary proposal on some ethical values and attitudes which should be shared by software developers aiming to achieve moral excellence in their professional performance:

First, awareness and acknowledgement of the moral dimension of software development.

Second, a general attitude of responsibility and preliminary investigation about implicit and declared intentions before use of any piece of already existing software.

Third, intentions should be declared as non-functional requirements for any new software. This will help to assume the moral consequences of choices and decisions made in professional contexts, mainly if those consequences are harmful, discriminatory, or lead to any unexpected misuses of software.

References

Gamez, P., Shank, D.B., Arnold, C. et al. (2020) Artificial virtue: the machine question and perceptions of moral character in artificial moral agents. AI & Soc 35, 795–809. https://doi.org/10.1007/s00146-020-00977-1
Hagendorff, T. (2020) The Ethics of AI Ethics: An Evaluation of Guidelines. Minds & Machines 30, 99–120. https://doi.org/10.1007/s11023-020-09517-8
Verbeek, PP. (2014). Some Misunderstandings About the Moral Significance of Technology. In: Kroes P., Verbeek PP. (eds) The Moral Status of Technical Artefacts. Philosophy of Engineering and Technology, vol 17. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-7914-3_7

Tactics & Rhetorics. Images of calculation models in architectural design competitions

Tilke Devriese

The presentation will discuss the agency of printouts of calculation models for a structure in competition proposals for architectural projects.

'Integrated design' as a quality marker has become an undisputed commonplace in building industries. This can be obtained through successful multidisciplinary collaboration. Architectural competitions are a typical setting in which this strive for multidisciplinarity and integration comes to the fore. The common stipulation that a proposal has to be formulated by a team, covering all necessary expertise, largely facilitates multidisciplinary discussions in early design phases. Although the relevance of multidisciplinary design in early design phases is beyond dispute, the contingencies of this (new) reality ask for reflection.

Often, technical aspects are extensively elaborated upon in competition output. Among others, images of calculation models are commonly used to illustrate this content. These images are design documents such as plans, sketches, scale models etcetera, all crucial design and communication tools for both the architect and the structural engineer. These documents contain models for the projected reality of the project. However, because of the different epistemic context from which team members operate, the understanding of these models does not fully coincide.

'(Models) are never handled or attributed meaning in isolation. They are part of a broad system of knowledge, composed of theories, material and mathematical artefacts, and interpretations.' (Loukissas 2008, p. 46)

The thesis is that the input of structural engineers in a competition proposal should as well be read in the light of the rhetoric used to try to win a competition. The images become active due to these differences in the understanding of the documents. Technical aspects, of which structural verification is one example, typically need concrete input. This information often is still open in early design phases, characteristic for a competition proposal. For professionals from the specific domain it is obvious that the ever present 'technical' images in competition proposals remain meaningless in terms of technical verification. But these images do demonstrate an early started multidisciplinary process. Moreover, a proposal that looks well studied and detailed is convincing and reassuring for a client.

Readers, in this case both client and architect, can only draw on prevalent preconceptions for their interpretation of the images. In this case, the paradigmatic understanding of the objectivity of the engineer's work in combination with the visual illiteracy to understand calculation models largely influences the impact of the images. This as well opens manoeuvrability for the engineer to further steer the design process.

My interest in the topic has grown since the day I was asked to 'quickly' shoot some 3D FEM-calculation graphs to illustrate a competition proposal, while working as an engineer. I hope the presentation will facilitate a fruitful discussion on theoretical concepts to frame this 'phenomenon' of calculation printouts in competition bundles.

References

Loukissas, Yanni. 2008. "Conceptions of design in a culture of simulation." Doctor of Philosophy in Architecture: Design and Computation, Department of Architecture Massachusetts Institute of Technology.

Weaponized artificial intelligence (AI), normative and culture psychology: Methodological contributions to policies on emerging technologies

Rockwell F. Clancy, Ingvild Bode & Qin Zhu

Our presentation outlines the motivations for and nature of a project exploring the development of global norms surrounding the use of weaponized AI. The motivations for this project consist in the need to develop effective policies surrounding the use of AI in weapons systems, which adequately account for the nature of normative thoughts and behaviors. The nature of this project consists in developing standardized measures of AI ethics and weaponized AI, coupled with questions exploring the development of norms, validated among different national and international stakeholders.

Motivations

The use of AI in weapons systems raises global security concerns and ethical questions, calling for the effective formulation and implementation of policies. To date, such work has tended to be theoretical and normative in nature, consisting in critical policy analyses and ethical considerations, carried out by philosophers, legal scholars, and political scientists. Little work has addressed what different stakeholders – from politicians to the public – know and think about weaponized AI. This information is necessary to effectively develop and implement policies, given the nature of normative thoughts and behaviors.

Laws have been the focus of work on AI in weapons systems in political science and international relations, for instance, how such technologies would (not) fit into existing regulatory frameworks. However, laws are only effective when they are supported by informal norms and corresponding institutions – for example, implicit social expectations regarding what we think others are doing (descriptive norms) or what we think they think we should be doing (social norms) – and culture affects these dynamics. Effective international policies on emerging technologies depend on better understanding these dynamics.

Methods

To begin to address these challenges, it would be necessary to understand not only what but also how different stakeholders think about weaponized AI. To do so, our project consists in developing instruments and carrying out research to capture explicit and implicit views among different stakeholders regarding (1) the development and use of AI in general and within weapons systems specifically and (2) how these are related to norms. Measures used to study ethics – such as the Defining Issues Test and Moral Foundations Questionnaire – will be adapted to AI ethics. These measures will assess the importance attached to different kinds of AI ethical concerns. Factors such as national, disciplinary, and professional background will be treated as input variables, helping to identify significant differences between stakeholder types. To understand the sources of these differences, additional measures will be used to explore norms.

Norms about weaponized AI cannot be understood in terms of the view of policymakers alone. Rather, they are affected by expectations regarding the views and behaviors of diverse stakeholders. To explore these dynamics, questions related to the development and use of AI will be coupled with ones to understand norms and the influences of different groups within reference networks. These will be administered and validated among samples from different national and international stakeholder groups.

Assessing student progress in engineering ethics education

Martin Sand, Michael Klenk & Andrea Gammon

Engineering ethics education is proliferating in universities (cf. Mitcham and Englehardt 2019). But measures to assess its success in terms of student progress are highly controversial, and reported effects are of varying magnitude. Success is most frequently measured by reactive measures (e.g. student satisfaction), which does not allow for reliable inferences about student's progress (Steele et al. 2016).

Moral reasoning is another popular measure of student progress in engineering ethics, measured by general measures of moral development such as the Defining Issues Test (DIT2) or domain specific approaches like the Engineering and Science Issues Test (Borenstein et al. 2010).

However, findings about the effects of engineering ethics courses on moral reasoning (as measured e.g. by the DIT2) are conflicting, and positive findings reveal modest effect sizes at best (Antes et al. 2009; Drake et al. 2005; Borenstein et al. 2010; Harris et al. 1996). Moreover, there are general concerns with taking a test of moral development in general (even if it is a valid measure of moral development) as a measure of the success of an engineering ethics course. A tailored and specific alternative approach to reactive measures and moral reasoning has been piloted by Davis and Feinerman (2012). However, their findings are based on a small sample size, and they do not include a control. In light of the general scientific desirability of replication, a confirmation of their methods and results is in order.

Moreover, research on student progress in engineering ethics has thus far set aside central insights from pedagogical psychology about important moderators of student success. Notably, there are robust findings about the positive effect of competence beliefs (roughly, beliefs about one's ability to do well in a certain subject) and control beliefs (roughly, beliefs about one's ability to take control of one's actions and one's situation) on student success (Trautwein et al. 2009; Wigfield and Eccles 2000; Wigfield and Eccles 1994). It is so far unclear how engineering ethics students evaluate their own competence in ethics and how far they believe that they are in control. Therefore, the present study tested a tailored measure of student progress in ethics and explored the effects of competence beliefs, control beliefs, and student expectations on student progress in engineering ethics. We used a quasi-experimental pre-test / post-test design. A questionnaire was administered online to students in a bachelor and master's level course in the academic years 2020/21 and 2022. We also evaluate a related potential moderator of student success that has so far not been systematically studied: student expectation. Own experience and the experience from other engineering ethics educators (Newberry 2004) suggest that students are often unsatisfied with their learning experience. Often, that dissatisfaction is related to learning success (Holsapple et al. 2011).

Our presentation will make two contributions. First, we will present the findings of the quasi experiment indicating whether students showed progress with regards to engineering ethics. We will then evaluate our findings critically vis-à-vis the validity of our proposed measure and a philosophically sound concept of moral development.

Translating ethical principles into computer codes: Experts' opinions on why (not) and how to do it

Franziska Poszler, Edy Portmann & Christoph Lütge

"[A]utonomous machines, i.e. machines equipped with automated decisions functions, are intended to be put in contexts where computed decisions have to be guided by ethical considerations" (Bonnemains et al., 2018; pp. 3-4). For example, algorithmic advisory systems are developed that aim to assist medical professionals in critical treatment decisions or autonomous vehicles have to decide which traffic participant to target when confronted with imminent crashes (Bonnemains et al., 2018). Therefore, it becomes critical to determine how to program such machines to behave in an ethically correct manner (Bringsjord et al., 2006). In this regard, scholars have suggested computational ethics as a promising frontier that aims to translate ethical principles into computer codes to ensure that the corresponding technology is both compliant with an ethical system and functionally dynamic to operate in the real world (Segun, 2021). Nevertheless, many outstanding questions concerning (long-term) consequences and the concrete implementation of computational ethics still prevail (Awad et al., 2022). Therefore, this study aims to contribute to this debate by addressing the following research questions:
1. What are arguments for and against computational ethics?
2. How could and should we realize computational ethics?
To shed light on these questions, this study draws on semi-structured interviews with eleven experts from the field of philosophy, AI and cognitive sciences. The interviews were audio-recorded, transcribed verbatim and coded manually using the MAXQDA software. In this analysis, we adopted an inductive coding methodology (Gioia et al., 2013) to identify themes as they emerged during data collection. Preliminary results suggest that indicated supporting and opposing arguments concerning computational ethics can be clustered into pragmatic/practical (e.g., technology companies' acknowledgement of the need to design for values), societal (e.g., manifestation and reinforcement of human biases) and epistemic (e.g., missing ground-truth about which ethical principles should be consulted) reasons. Furthermore, stated recommendations of how to realize computational ethics are recapitulated (e.g., technical solutions such as machine learning, necessary restrictions in the design of corresponding technologies such as their limited operation space, relevant alternatives/complementary activities such as the education of programmers). Lastly, indicated open questions are summarized that need to be addressed in the future (e.g., which ethical principles to use as reference points?, how and with whom do we agree on that?).

References

Awad, E., Levine, S., Anderson, M., Anderson, S. L., Conitzer, V., Crockett, M. J., ... & Tenenbaum, J. B. (2022). Computational ethics. Trends in Cognitive Sciences.
Bonnemains, V., Saurel, C., & Tessier, C. (2018). Embedded ethics: some technical and ethical challenges. Ethics and Information Technology, 20(1), 41-58.
Bringsjord, S., Arkoudas, K., & Bello, P. (2006). Toward a general logicist methodology for engineering ethically correct robots. IEEE Intelligent Systems, 21(4), 38-44.
Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking qualitative rigor in inductive research: Notes on the Gioia methodology. Organizational research methods, 16(1), 15-31.
Segun, S. T. (2021). From machine ethics to computational ethics. AI & SOCIETY, 36(1), 263-276.

Technological Optimism vs. Scepticism: Agroecology as a Case Study

Alik Pelman

As indicated by the widely accepted Kaya identity (Kaya et al, 1997), the main drivers of environmental burdens (and in particular of carbon emissions) are: 1. Population; 2. Income, i.e., expenditure per person; 3. Energy intensity, i.e., energy consumed per money spent; and, 4. Carbon intensity, i.e., carbon emissions per each energy unit (the latter two together amount to 'technological eco-efficiency'). Multiplying these four parameters results in the total carbon emissions caused by a given population (e.g., humanity as a whole).

Between 1965 – when environmental awareness was already on the rise – and 2018, total emissions increased by +224%. This is mainly due to the first two drivers: population growth (+130%) and income growth (+155.6%), whereas energy intensity (-45%) and carbon intensity (-13%) managed to slightly slow this trend down. (Source: Our World in Data based on Global Carbon Project; UN; BP; World Bank; Maddison Project.)

Nevertheless, policy makers overall avoid meddling with population size (family planning) or consumption (economic growth), hence virtually all of humanity's efforts are concentrated on technological solutions. But as the above numbers show, current technological efforts are extremely far from being able to meet the challenge put by population growth and economic growth – which are mostly considered untouchable.

Most technological efforts to address the environmental crisis take the form of advanced technologies, e.g., nanotechnology, biorobotics and the like. Technological sceptics mistrust technology as a means to solve the environmental crisis due to its slow improvement rate, and due to fear of the high risks involved in such large-scale interfering technologies (e.g., sulphate aerosol injection, bio-energy and carbon capture and storage (BECCS), afforestation, and the like). The purpose of the present study is to propose another route of technological eco-efficiency, which is based on appropriate technologies rather than on techno-science.

We take as our study case a small (750 m2) low-input subsistence Mediterranean agroecological farm in a developed nation, that is based on intercropping and annual crop rotation. The farm provides one individual, the proprietor, with nutritional self-sufficiency (adequate intake of an array of macro- and micro-nutrients) with limited labor, no synthetic fertilizers or herbicides, and zero waste, effectively closing a full farm-table-farm cycle. We find that the agroecological farm outperforms industrial farming in terms of both lower environmental burdens, across all examined environmental metrics (70% lower on average) per kg produce, and higher nutritional score (40% higher on average). Per ha, the environmental lopsidedness was even higher (82% lower than industrial farming on average), but lower nutritional score (18% lower on average). However, when considering total land used by industrial farming, i.e., including all non-agricultural land, per ha farmland, the nutritional score of the agroecological farm considerably outperforms that of industrial farming (140% higher on average).

Agroecology thus presents answers to both worries of the technological sceptics, namely, it can be implemented immediately with dramatic results, and the technologies it uses are of very low risk. As such, agroecology may point at a new paradigm of technological optimism that is based on appropriate technologies rather than on techno-science.

Towards A Philosophy of Bioengineering and Geoengineering: Risk Ethics for a Climate Change Technofix

Joshua Wodak

In light of the hysteresis and acceleration of the climate crisis, climate overshoot has only recently been acknowledged as inevitable. As the IPCC belatedly reports, current pledges are not even remotely on track to limit global warming to 1.5°C above pre-industrial levels (Anderson 2015, IPCC 2018). Further, no amount of future emissions reductions can suffice to avert climate overshoot. Such is the rise of interest, born of utter desperation, in a climate change technofix.

Hence, this presentation explores the intersection of engineering and philosophy, using the framework of applied ethics to critically analyse the proposition that a climate change technofix – namely Negative Emission Technologies (NETs) – is the only potentially efficacious means to avert runaway climate change (Carton 2020, Reynolds 2015).

However, not only is the efficacy of NETs to reduce sufficient greenhouse gas concentrations highly dubious, but any such technofix requires gambling on a host of unknown unknowns – namely, the inexorable complexity of the Earth System, coupled with planetary-scale interventions in the crisis. Therein, I draw on ethics, social and political philosophy to put forth a critique of how normative ethics remains anchored in rigid positions of anachronistic risk aversion, given how any attempted climate technofix entails unprecedented realms of risk and uncertainty. Drawing on the work of applied ethicists Clive Hamilton (2010 and 2017), Ronald Sandler (2012 and 2015), and Christopher Preston (2013 and 2018), I critically engage with the risk ethics of imminent climate overshoot, in relation to the interventionist gambles proposed by NETs through Synthetic Biology and Climate Engineering. Given the scale of the unknown unknowns unleashed by the Anthropocene, I present gambling as the most apt analogy for both the absurdity (and denied imminence) of the existential predicament, as well as the sheer improbability that any technofix can be invented in a sufficiently short time and implemented on a sufficiently large scale. Therein, the presentation contemplates the unthinkable questions that our current situation demands we ask, and perhaps even try to answer.

References

Carton, W. 2020. 'Carbon Unicorns and Fossil Futures: Whose Emission Reduction Pathways is the IPCC Performing?' In J. Sapinski, H. Buck and A. Malm (eds), Has it Come to This? The Promises and Perils of Geoengineering on the Brink, pp. 34–49. London: Rutgers University Press.
Hamilton, Clive. 2010. 'The Return of Dr Strangelove. The politics of climate engineering as a response to global warming'. In Rosalie Bertell (ed.), Kriegswaffe planet erde, pp. 468–490. Gelnhausen-Roth: Fischer-Verlag.
Hamilton, Clive. 2017. Defiant Earth: The Fate of Humans in the Anthropocene. Allen & Unwin.
Preston, Christopher, ed. Engineering the Climate: The Ethics of Solar Radiation Management. Lanham: Lexington Books, 2013.
Preston, Christopher. The Synthetic Age: Outdesigning Evolution, Resurrecting Species, and Reengineering Our World. Cambridge: MIT Press, 2018.
Reynolds, Jesse. 2015. 'A Critical Examination of the Climate Engineering Moral Hazard and Risk Compensation Concern'. The Anthropocene Review 2(2): 174–191.
Reynolds, Jesse. 2021. 'Earth System Interventions as Technologies of the Anthropocene'. Environmental Innovation and Societal Transitions 40: 132-146.
Sandler, Ronald. The Ethics of Species: An Introduction. Cambridge: Cambridge University Press, 2012.
Sandler, Ronald, and John Basl, eds. Designer Biology: The Ethics of Intensively Engineering Biological and Ecological Systems. Lanham: Lexington, 2015.

Towards a global ethics of artificial intelligence (AI): Ensuring AI ethics is culturally responsive and psychologically realist

Rockwell F. Clancy, Qin Zhu, Louis Hickman, Subhabrata Majumdar, David McGraw & Andrew Katz

Our presentation outlines the motivations for and nature of a culturally responsive, psychologically realist approach to global AI ethics. The motivations for this project are the fact that (1) AI is global in nature, and (2) traditional approaches to AI ethics are potentially ineffective. The nature of this approach consists in developing tools for and conducting research regarding what peoples from different cultures, countries, disciplinary, and professional backgrounds think about AI and ethics.

Motivations

First, AI is global in nature, as it affects billions of lives across countries and cultures. Seemingly technical decisions in AI design have ethical import, since they can perpetuate and reinforce injustices and harms. These dynamics highlight the importance of policies for ethical AI. Although companies and academics worldwide have worked on the development of AI, market concentration has occurred in only a few regulatory jurisdictions, such that AI policies have represented a narrow set of ethical concerns. Intergovernmental and multistakeholder initiatives have begun to address these issues, but it is unclear if contributing individuals represent the concerns of those affected by AI. Hence, a global approach to AI ethics must be culturally responsive.

Second, traditional approaches to AI ethics are potentially ineffective, since they have tended to be normative in nature and based on mistaken assumptions regarding how people think about issues of right and wrong. Such approaches have consisted in outlining values or principles regarding what should or should not be done, focusing on trust, transparency, fairness, and rights. These have served as the basis for policies on and training in ethical AI. However, neither ethical judgments nor behaviors are based primarily on ethical reasoning. As a result, studies have found that AI ethics initiatives to date rarely have an impact on decision-making in the field. Hence, a global approach to AI ethics must be psychologically realist.

Nature

The nature of this project involves developing tools to study what people think about AI and ethics. This consists in four related steps that build on each other: (1) using natural language processing algorithms (NLP) on documents from international academe, governments, industries, and inter- and non-governmental organizations, to identify AI ethical perspectives recognized as important by those responsible for developing AI and AI-related policies; (2) conducting semi-structured interviews with samples of individuals from the groups and organizations in step (1), as well as with ones from outside these groups – including members of the public, with varying degrees of education and technical understanding/knowhow – to identify AI ethical perspectives that have not been recognized as important by those responsible for developing AI and AI-related policies; (3) findings from steps (1) and (2) will be used to develop a standardized survey to understand not only what but also how people think the ways they do about AI and ethics; (4) this instrument will be assessed in terms of its ability to predict perspectives on the use of AI-related technologies.

Introducing the concept of hidden morality in energy justice

Nynke van Uffelen

Technological systems, including energy systems, shape to a large extend our sociotechnical world. Therefore, there is a moral imperative that people living in that world are able to influence its course. As such, it is important to understand what citizens conceive as just or unjust related to the energy transition. In this respect, scholars in design for value and responsible research and innovation increasingly call for more inclusive participation procedures in engineering decision-making. Similarly, the energy justice scholarship aims to understand claims of injustice, voiced for example in energy conflicts, by categorising them into tenets. Such calls for more public participation trickled down to energy policies: the Dutch government, for example, prescribes local participation processes in the energy transition (Klimaatakkoord, 2019).

Current methods for detecting injustices rely on explicit articulations of beliefs by citizens in official participatory settings or energy conflicts. However, it is implausible that all injustices manifest within these contexts. Many municipalities in the Netherlands struggle to organise inclusive participation trajectories that involve a truly diverse publics. As such, to make energy systems more just it is important to explore which mechanisms prevent certain injustices from surfacing, in other words, what the problem is.

This study introduces a concept that is helpful to understand why injustices might remain unseen and unaddressed, namely the problem of hidden morality (Honneth, 1995a). Its proponent, the philosopher Axel Honneth, theorises that there are several steps between the occurrence of injustices and social change, namely: (1) injustices are experienced as “negative emotional reactions” (Honneth, 1995b, p. 136); (2) injustices are expressed in the form of claims of injustice; (3) people collectively organise themselves and engage in collective action; (4) claims are taken up in the public discourse; (5) claims are reformulated in a positive manner; and (6) actual social change. Between each of these steps, different obstacles can arise.

In this study, the problem of hidden morality is proposed as a promising avenue for future research on detecting and understanding energy injustices, for two reasons. First, the concept is a helpful alternative to existing frameworks (such as distinguishing tenets of justice) for detecting energy injustices, because it allows for distinguishing mechanisms that prevent injustices to surface. Second, the problem of hidden morality prescribes methodological innovation, as uncovering the different barriers between different steps requires multiple research methods from diverse academic disciplines. In sum, understanding which and why injustices remain hidden is the first step towards making energy systems more just.

References

Honneth, A. (1995a). Moral Consciousness and Class Domination: Some Problems in the Analysis of Hidden Morality. In C. W. Wright (Ed.), The Fragmented World of the Social: Essays in Social and Political Philosophy. State University of New York Press.
Honneth, A. (1995b). The Struggle for Recognition: The Moral Grammar of Social Conflicts (J. Anderson (Ed.)). MIT Press.
Klimaatakkoord. (2019). https://www.rijksoverheid.nl/documenten/rapporten/2019/06/28/klimaatakkoord
McCauley, D., Heffron, R., Stephan, H., & Jenkins, K. (2013). Advancing Energy Justice: The Triumvirate of Tenets. International Energy Law Review, 32(3), 107–110.

Designing Automation with Care

Michael Wartmann

Manufacturing processes have increasingly become automated over the past decades. Starting out with small-scale hand-operated processes in the early times, current state of the art can in specific cases reach near autonomy and remote operation for large and complex processes. The drive for automation is twofold: Primarily, improving the productivity of the process by lowering cost and increasing throughput ever further involving advanced sensors, computer models, and algorithms. On the other hand, finding ways to enhance plant and operator safety, ergonomics, and general reliable operability of the respective process plant. This drive for automation has led for the common process operators, engineers, and managers to transition from their work being a craft with significant freedom of design to an increasingly passive and procedural professional frame with few but potentially severe exceptions during abnormal operation.

In this work, I discuss how technology relations in technology-mediated organizations often unknowingly and implicitly define the ethical perspective by the way automation technology is adopted. Realizing the causality dilemma of a human organization and its applied technology being tightly intertwined, I propose reevaluating the perspective on automation in industry. Automation requires balancing the concept of Heidegger's releasement in combination with critical thinking towards an ethical dialogue throughout all technology design phases. In this design approach, technological maturation runs in parallel to psychological maturation of the workforce and their leadership. Eventually, organizations can adopt concepts such as Nussbaum and Sen's capability approach helping them define their ethical perspective of care for workforce, customer, society, and nature leading to sustainable technology deployments.

These technology-mature organizations would understand the fundamental nature of human-technology relations and design for it through having creative capabilities of the connected individual expressed, understood, and mediated while realizing that technology is an extension of humanity and our ethics rather than a neutral tool.

Generic design Science and lived experiences

Eswaran Subrhamanian, Anne-Françoise Schmid, Ira Monarch, Muriel Mambrini-Doudet & Anne-Lise Dauphiné-Morer

We approach design science in interdisciplinary fashion. We include not only the sciences, both natural and human, but also the disciplines of philosophy, art, engineering and architecture. We also include and how these interact with the public and market realms. The collision of these disciplines can inspire the sciences to focus on what can and cannot be done as with constructor theory in the physical sciences and what is being or to be done as with design theory in the human sciences. We keep the word "science" to emphasize knowing as well as doing. We focus on design science in our presentation. For philosophy, we emphasize, generic epistemology, esthetics and ethics. Our view of the interdisciplinary science of design assembles evolving design knowledge including new design concepts through shared memory of the feedback relationships between what we call the theory of the design case and what we call the practice or pragmatics of the design case. It also provides a generic space for enabling design intentions (based on the theory of the design case) to become realized in design pragmatics via building collective intimacy among the design group (even for individuals to sustain intention and practice over time). Moreover, shared memory and collective intimacy are interrelated via fictional, virtual and futural operators and integrative objects, models and simulations. We provide various diagrams of the interrelationships among these ideas in several geometric figures, all of which are octahedrons. There is some controversy in our group as to how these octahedrons should be labeled and what we intend to express by using them. One proposal we are working on is that the octahedron configures from the outside what lived experience and emotions might be like on the inside of design projects. It is this lived experience on which the accumulation, composition and elaboration of insights for the assembly of shared memory and realization of intentions in practice through collective intimacy is based. We describe the experience of two members of our group concerning the effects of the octahedron on them and the implications of their experience for interpreting design theory and design science. We also discuss the many feedback/feedforward relationships leading to the elaboration of an intention and theory of the case for a project several members of the group are working on - a new of view of engineering education in post-colonial India. Riffing from these two vignettes, we show how innovative designs compose new hyper-compatible concepts, but also how local projects lead to the refinement of a generic theory of design.

The moral authority of ChatGPT

Sebastian Krügel, Andreas Ostermaier & Matthias Uhl

ChatGPT, OpenAI's cutting-edge AI-powered chatbot, is a brilliant and engaging conversationalist, which solves exams, writes poetry, and creates code. It also searches information, answers questions, and gives advice. As users may rely on ChatGPT's advice for consequential decisions, important ethical questions arise. Indeed, people often hold contradictory moral beliefs, and with consistent moral advice, ChatGPT might improve users' judgment and decisions. However, is ChatGPT a morally consistent advisor? Does its advice influence users' moral judgment? And does this influence depend on whether users know that they are advised by a chatbot? We ran a two-stage experiment to answer these questions. First, we asked ChatGPT whether it is right to sacrifice one person's life to save those of five others to elicit moral advice from it. Second, we presented subjects with the trolley problem, which features this exact dilemma, along with ChatGPT's answer. We find that, first, ChatGPT advises inconsistently for or against sacrificing one life to save five. Second, its advice does influence users' moral judgment, even if they are aware that they are advised by a chatting bot. Third, users underestimate the influence of ChatGPT's advice on their judgment.

Scenarios for the development of SRM between 2030-2050 with particular reference to the interaction between SRM and Climate Change and Global Catastrophic and Existential Risk

Gideon Futerman

This session will present some of the results from a ParEvo scenario exercise run on plausible futures related to SRM development, governance and deployment between 2030-2050, with particular reference to its interaction with Global Catastrophic and Existential Risk (GCR and XRisk). The session will introduce the ParEvo method, lay out some of the futures generated, and discuss the evaluation of the exercise, as well as some of the discussions presented at a workshop to evaluate the exercise in Cambridge at the end of March.

As a potentially highly impactful emerging socio-technical system, understanding plausible (if speculative) futures of SRM will be exceptionally useful. However, traditional scenario generation methods are often relatively constrained to physical characteristics of deployment, whereas ParEvo allows a qualitative and highly exploratory exploration of physical, technological, social and political characteristics of the technology, all which are especially relevant to the development of a socio-technical system like SRM and its interaction with extreme risk (GCR and XRisk). The ways in which these features were relevant to the futures will be explored. Moreover, the method generated a variety of governance frameworks for assessments, and stress tested a number of these in useful and informative ways, so the utility of the exercise in informing SRM governance will also be discussed. The inherent transdisciplinarity of this foresight method also lends itself to generating informative scenarios about GCR/XRisk specifically, and so this will also be explained, and the utility of this method to evaluating SRM and its relation to XRisk will further be discussed.

The scenarios were generated by 15 researchers in or around SRM, and the utility of the process lay in part in the information it provides about conceptions of futures by members of the field. The evaluation of the exercise, where common features of futures, participants impressions of the futures' plausibility and other features of the futures, attitudes towards SRM expressed in the futures and more will also be discussed.
This will all be brought together to discuss the contribution of SRM and Climate Change to GCR/XRisk, and how scenario generation such as this can be useful at reducing risk.

Automation and autonomy: How sociotechnical systems disrupt the counterfactual status of autonomy

Udo Pesch

Moral autonomy is a key value in modern ethics (Kant, 2017). Often this value is taken as a given, as an intrinsic human characteristic from which ethical postulates are derived. Subsequently, institutions and technologies are designed in such a way that they align with these postulates. In this paper, I see autonomy as a counterfactual trait, in the sense that it can become real in empirical reality (cf. Habermas, 2022). A counterfactual approach to values seems productive in case of autonomy. It is not the question, whether autonomy is an 'ought' or an 'is', but the question is what it can be. From this counterfactual starting point, the design of institutions and technologies should not so much be based on the assumption of autonomy, but on their ability to nurture autonomy as a potential. In other words, the question needs to be asked what it would mean for the design of institutions and technologies, if persons are autonomous.

In other words, institutions and technologies can facilitate people to become rational, so that the moral ideal of autonomy can become substantiated (Pesch, 2022). Acknowledging the retrospective reconstruction, the legal system can be seen as representative for the way that modern institutions function: they are social contexts in which individuals that have transgressed existing laws are asked to explain what motivated them make certain choices, so that they are given the opportunity to attune decisions, actions and consequences.

This counterfactual approach to autonomy allows a refined assessment of new technologies and institutions, especially within the context of the increasing dissemination of automated systems throughout society. It is typical that these sociotechnical systems not so much assume individuals as autonomous beings, but that they tend reproduce a behaviouristic inclination which reduces the capacity of individuals to choose for themselves who they want to be. Institutional developments that do not so much nurture autonomy as a potential, but come to confine it.

References

Habermas, J. (2022). Ein neuer Strukturwandel der Öffentlichkeit und die deliberative Politik. Frankfurt: Suhrkamp.
Kant, I. (2017). Kant: The metaphysics of morals: Cambridge University Press.
Pesch, U. (2022). Values as Hypotheses and Messy Institutions: What Ethicists Can Learn from the COVID-19 Crisis. In M. J. Dennis, G. Ishmaev, S. Umbrello, & J. van den Hoven (Eds.), Values for a Post-Pandemic Future (pp. 129-144). Cham: Springer International Publishing.

Ethics and sustainability concepts for the factory of the future: what should we teach and how?

Emmanuel Caillaud

The factory of the future will have to deal with several transitions: ecological, social, climate, digital and geopolitical.
The factory of the future must obviously be sustainable.
The evolution of industries raises a series of ethical questions.
How can we train engineers and managers in these issues? How can we train them to deal with ethical dilemmas?

A first experiment was carried out at the Conservatoire National des Arts et Métiers Paris (CNAM) and it is a question of sharing what was done but also the questions raised.
The proposed course is aimed at Master's students in the fields of quality, environmental safety and corporate social responsibility. It has been proposed to a cohort of students in French and a second time to international students (Master in Management Sustainable Development and Quality Management).

The course was structured as follows:
- introduction to ethics
- main concepts
- main works
- case studies dealing with ethical dilemmas for managers and engineers involving several transitions.

Elements that raise questions:
- what concepts to teach?
- which reference books to consider?
- how far can a non-philosopher teach ethics?

References

Freeman, R. Edward. 2010. Strategic Management: A Stakeholder Approach. Cambridge University Press.
Jonas, Hans. 1985. The Imperative of Responsibility: In Search of an Ethics for the Technological Age. New edition. Chicago (Ill.): University of Chicago Press.
Newberry, Byron, Katherine Austin, William Lawson, Greta Gorsuch, and Thomas Darwin. 2011. "Acclimating International Graduate Students to Professional Engineering Ethics." Science & Engineering Ethics 17 (1): 171–94. https://doi.org/10.1007/s11948-009-9178-6.
Trentesaux, Damien, and Emmanuel Caillaud. 2020. "Ethical Stakes of Industry 4.0." IFAC-PapersOnLine 53 (2): 17002–7. https://doi.org/10.1016/j.ifacol.2020.12.1486.
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.