

# CEPE 2023: Human-AI Interaction and the Future

VIRTUAL CONFERENCE: MAY 24–25



**INSEIT**

International Society for Ethics and Information Technology

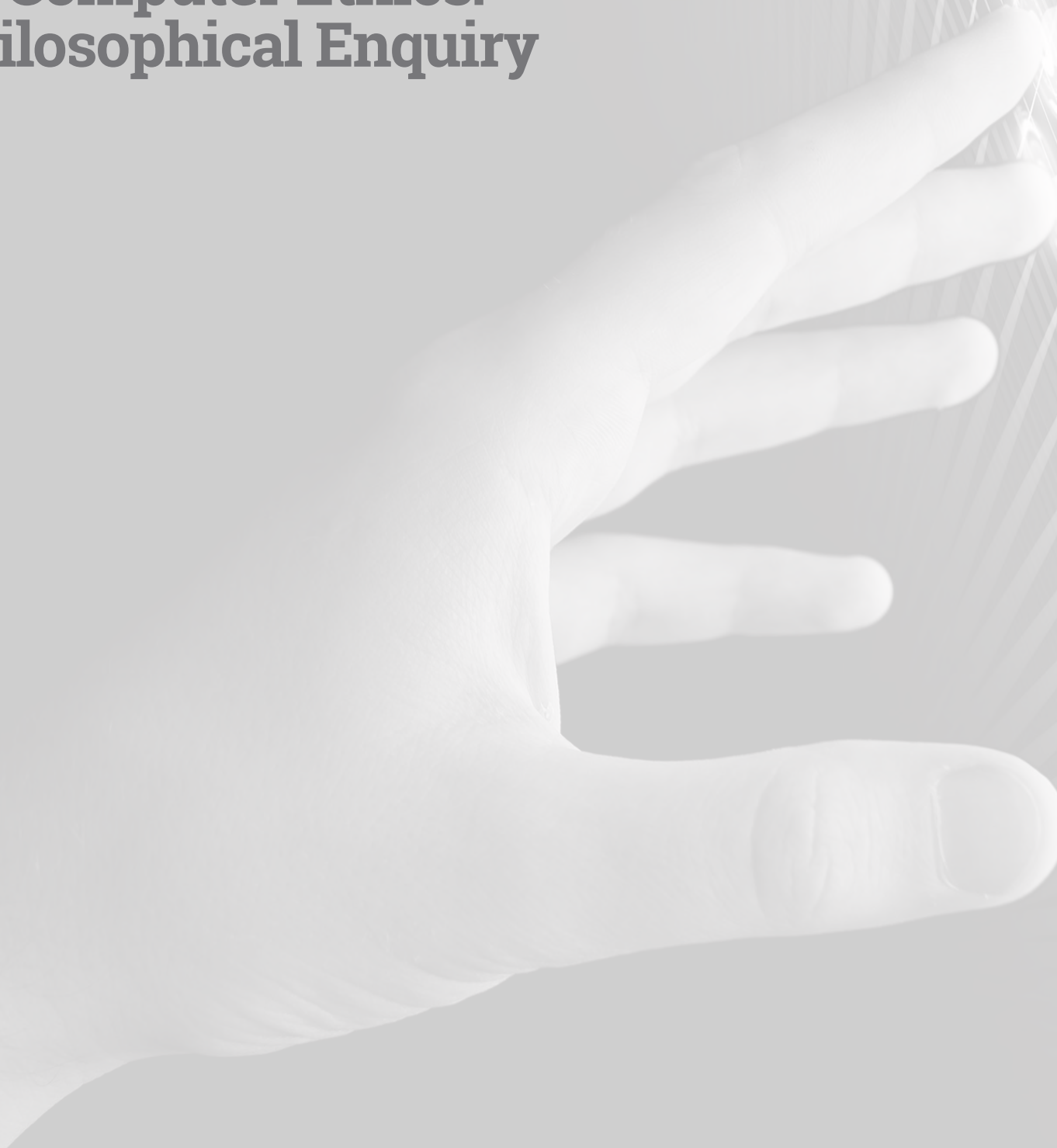
**ILLINOIS TECH**

Center for the Study of  
Ethics in the Professions

Center for Cyber Security  
and Forensics Education

# **CEPE 2023**

**International Conference  
on Computer Ethics:  
Philosophical Enquiry**



# Schedule at a Glance

Wednesday, May 24	Thursday, May 25
7:00 a.m. CDT/1:00 p.m. CET Welcome	8 a.m. CEST/4 p.m. KST 6 p.m. AEST/8 p.m. NZST Session 4
7:15 a.m. CDT/1:15 p.m. CET Session 1	7 a.m. CDT/1 p.m. CET Session 5
9:15 a.m. CDT/3:15 p.m. CET Break	9 a.m. CDT/ 3 p.m. CET Break
9:30 a.m. CDT/3:30 p.m. CET Session 2	9:15 a.m. CDT/3:15 p.m. CET Session 6
11:00 a.m. CDT/5:00 p.m. CET Break	11:15 a.m. CDT/5:15 p.m. CET Break
11:15 a.m. CDT/5:15 p.m. CET Session 3	1130 a.m. CDT/5:30 p.m. CET Session 7

## Day 1 Meeting Links

### Room 1:

<https://iit-edu.zoom.us/j/85770308748?pwd=VFQrTVhGV2xadmhFQklpbzZMDlwZz09>

### Room 2:

<https://iit-edu.zoom.us/j/88159904862?pwd=MUJsYzFlTXhkNGRYYWxoaDZ6UStzQT09>

## Day 2 Meeting Links

### Room 1:

<https://iit-edu.zoom.us/j/89105975721?pwd=RUZsVGJFN1laTkhMYzFOUnU3S0pVUT09>

### Room 2:

<https://iit-edu.zoom.us/j/88159904862?pwd=MUJsYzFlTXhkNGRYYWxoaDZ6UStzQT09>

*Times listed are in Central Standard Time and Central European Time*

# Day 1—May 24

7 a.m. CDT/ 1 p.m.  
CET

## Welcome

Elisabeth Hildt

*Center for the Study of Ethics in the Professions, Illinois Institute of Technology*

## Session 1

7 :15–9:15 a.m. CDT  
1:15–3:15 p.m. CET

### Room 1

#### AI and Law

<https://journals.library.iit.edu/index.php/CEPE2023/article/view/235>

**Chair:** Maria Bottis

#### **Freeing Digital Images at Last? The Interplay of Copyright, Public Domain, New Technologies, and Ethics in Museum Reproductions of Visual Art**

Maria Bottis, *Ionian University, Greece*

The presentation deals with the laws and ethics on museum digital reproductions of visual art. Copyright protects works which are original but very often, museums collect and exhibit works which have fallen into the public domain and also, are considered legally as ‘monuments’, protected by strict cultural property laws and not only copyright law. In Europe, Directive 790/2019 has freed visual art works when the copyright term has ended, but the implementation in the European member states has not always dignified this mandate. The presentation aims at offering a balanced approach towards the ends of preservation and dissemination of digital reproductions of works in museums which are part of the public domain, but also, monuments. Museum ‘copyfraud’ is explained, as it harms free speech and the dissemination of cultural objects. The Boticelli Uffici lawsuit against Gaultier is examined. The presentation will then present solutions towards the protection of public domain in line with the invaluable mission of museums worldwide.

#### **European or Universal? The European Declaration of Digital Rights in a Global Context**

<https://journals.library.iit.edu/index.php/CEPE2023/article/view/250>

Maksymilian Kuźmicz, *Stockholm University, Sweden*

Technological innovations create great opportunities, but they come with big challenges too. While they allow people to be more autonomous and enjoy their lives, they may make people feel more dependent on technologies that they do not necessarily understand completely. This might create an atmosphere of anxiety and insecurity for end-users. Balancing various interests and protecting the rights of all stakeholders represents a legal challenge. One solution might be to draft a new catalogue of fundamental rights. In the EU, the European Declaration of Digital Rights is being proceeded. But can it be the first step toward establishing Digital Rights on the global level? In order to answer this question, we shed light on some fundamental philosophical and legal principles to evaluate the possibility of a Universal Digital Rights Declaration. Similar, to previous research on Universal Human Rights, this approach is intended to foster a discussion about the importance of global norms when it comes to individuals’ rights during technological development and implementation. This is the first step to supporting technological innovation and protecting the people who are affected by it.

#### **Trusting the Untrustworthy- A New Dimension to Situating Trust in Artificial Agents**

Omkar Chattar, *Indraprastha Institute of Information Technology, Delhi, India*

The problem of trust in AI is a convoluted one. Trust in artificial agents cannot be based only on reliability, efficiency, and precision, given that these agents are actively working for more than one user simultaneously. In this essay, we will see how current trust models exist within a framework of multiple users at different levels of abstraction and telos of the same artificial agent. We use case studies of various domains where artificial agents work in different telos like social media, streaming services, healthcare devices, autopilot systems, and car management systems, etc. Using Bruno



Latour's actor-network theory to analyze the systems and how human actors at different nodes change the nature of the Network itself and how the problem of trust changes from trust to the trustworthiness of actors in the Network, how the fidelity of Artificial agents towards their different users leads to a lack of trust. And how ownership of artificial agents is a central factor in determining their trustworthiness and building a sustainable human-technology trust relationship.

## **Room 2**

### **AI in Society**

**Chair:** Kelly Laas

#### **Ethical Deepfakes? The Values Held by Deepfake Developers and Creators, and their Governance Potential**

Maria Pawelec, *University of Tuebingen, Germany*

Deepfakes are proliferating in all areas of digital life, enabling benefits e.g., for digital self-expression or commercial gain, but also posing risks when employed for political manipulation, cybercrime, and non-consensual deepfake porn. Debates on deepfake governance are thus gaining momentum. These debates and related research largely neglect deepfakes' sources, namely the developers of the underlying artificial intelligence technologies, and the creators of actual deepfakes. This is noteworthy, since the growing field of Artificial Intelligence (AI) ethics suggests that these actors' values influence which deepfakes are and can be developed. This paper seeks to help fill this research gap by empirically investigating and comparing the values held by the various actors involved in developing and creating deepfakes, i.e., academics, open-source collaborators, and machine learning professionals working in large companies and start-ups. It then relates these values to broader debates on open-source norms and AI ethics. Lastly, the paper assesses whether the values identified (could) act as strong governance mechanisms curbing the creation of harmful deepfakes. In so doing, the paper contributes both to a more nuanced understanding of AI ethics in relation to deepfakes, where AI is not employed for algorithmic decision-making, and to empirically informing deepfake governance.

#### **Artificial Intelligence and Ethics: Ensuring technology for the Good of Society**

Anthony Novaes, *Universidade Presbiteriana Mackenzie, Brazil*

Artificial Intelligence (AI) has many definitions, amongst which, information systems which develop tasks similarly to human intelligence, and which have autonomy and are able to learn. AI can have positive impacts, such as their ability to predict the future, as consumer's behavior, which became of utmost importance after significant changes like the ones which took place in consequence of the Covid-19 pandemic. At the same time, many countries and regions start to regulate such matter, as United States, European Union and Brazil. Whilst the regulations vary in nature and complexity, the challenges to be faced are quite similar. Despite the brilliant opportunities which could be enabled by Artificial Intelligence, the bias and prejudice could be easily reproduced by the consequent solutions, reason why this issue needs to be properly tackled. In this sense, ethics play a key role and are fundamental for ensuring that AI serves society and fosters its improvement.

#### **Which Method for Engineering Concepts and Technologies?**

<https://journals.library.iit.edu/index.php/CEPE2023/article/view/256>

Irene Olivero, *Polytechnic University of Milan, Italy*

The introduction and increasing use of Artificial Intelligence in several areas of our lives has brought significant changes and challenges. AI is the common denominator of several so-called "socially disruptive technologies" (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 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 "re-shape" our socially disruptive technologies. As we do with concepts, asking ourselves what function these technologies 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. 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.

## **A Five-Step Ethical Decision-Making Model for Self-Driving Vehicles: Which (Ethical) Theories Could Guide the Process, and What Values Need Further Investigation?**

Franziska Poszler, Maximilian Geisslinger, and Christoph Lütge  
*Technical University of Munich, Germany*

By choosing a specific trajectory (especially in accident situations), self-driving vehicles (SDVs) will implicitly distribute risks among traffic participants and induce the determination of traffic victims. Acknowledging the normative significance of SDVs' programming, policymakers and scholars have conceptualized what constitutes ethical decision-making for SDVs. Based on these insights and requirements formulated in contemporary literature and policy drafts, this article proposes an explicit five-step ethical decision model for SDVs during hazardous situations. In particular, this model states a clear sequence of steps, indicates the guiding (ethical) theories that inform each step, and points out a list of values that need further investigation. This model, although not exhaustive and resolute, aims to contribute to the scholarly debate on computational ethics (especially in the field of autonomous driving) and serves practitioners in the automotive sector by providing a decision-making process for SDVs during hazard situations that approximates compliance with ethical theories, shared principles and policymakers' demands. In the future, assessing the actual impact, effectiveness and admissibility of implementing the here sketched theories, values and process requires an empirical evaluation and testing of the overall decision-making model.

**9:15–9:30 a.m. CDT**  
**3:15–3:30 p.m. CET**

### **Break**

**9:30–11:00 a.m.**  
**CDT 3:30–5:00 p.m.**  
**CET**

### **Session 2**

#### **Room 1**

#### **Ethical Decision-Making and Responsibility**

**Chair:** Franziska Poszler

#### **Responsibility Before Freedom: Closing the Responsibility Gaps for Autonomous Machines**

<https://journals.library.iit.edu/index.php/CEPE2023/article/view/266>

Shervin Mirzaeighazi and Jakob Stenseke, *Lund University, Sweden*

The introduction of autonomous machines (AMs) in human domains has raised challenging questions about the attribution of responsibility; referred to as the responsibility gap. In this paper, we address the gap by positing that entities cannot be granted the freedom of action—and be subject to blame or punishment—unless they can also recognise the same right for others. Since AMs—presently and in the foreseeable future—fail to meet this criterion, the user who utilizes an AM to pursue their goals can instead grant the machine their (the user's) right to act autonomously on their behalf. Thus, an AM's right to act freely hinges on the user's duty to recognise others' right to be free. Since responsibility is attributed before an entity is given the freedom to act, the responsibility gap never arises. We also consider a relevant objection: that holding users responsible for AM's actions will disincentivise manufacturers to produce safer machines. We show that this objection is misguided, as it wrongfully treats the responsibility gap as an excuse. Finally, we conclude by arguing that, if users are held responsible, it promotes a responsible use of AI while strengthening the incentives that motivates companies to make safer machines.

#### **From HHI to HRI: Which Facets of Ethical Decision-Making Should Inform a Robot?**

<https://journals.library.iit.edu/index.php/CEPE2023/article/view/252>

Jason Borenstein and Ronald Arkin, *Georgia Institute of Technology, United States*

Arthur Melo Cruz and Alan Wagner, *Penn State University, United States*

Over time, roboticists and others within computing communities have sought to create technology that seeks to emulate complex ranges of human-like behavior, potentially including the ability to participate in complicated conversations. Regardless of how sophisticated its functionality is, a robot should arguably be encoded with ethical decision-making parameters, especially if it is going to interact with or could potentially endanger a human being. Within this context, our NSF-funded research team is investigating some of the potential parameters that could inform a robot's ethical

decision-making processes when it interacts with humans. Whether insights derived from human-human interaction carry over to the ethical appropriateness of human-robot interaction is largely an open question. But our goal is to highlight some of the dimensions of ethical decision-making that occurs between humans, which warrant examination while the enterprise of encoding ethical robots proceeds

### **Human-AI interactions and the Moral Responsibility Gap. Re-Engineering Moral Responsibility Attribution in Low-Code/No-Code AI Platforms**

Samuela Marchiori, *Delft University of Technology, Netherlands (Not recorded)*

Low-code/no-code AI platforms allow virtually anyone with a computer and an internet connection to develop AI systems autonomously in a fast, easy, and inexpensive way, without the need for expert human supervision. This is concerning in that it delegates ethically charged development choices to individuals who may not have the necessary skill set to grasp their significance. This also contributes to exacerbating the severe responsibility gap that is frequently raised in the context of human-AI interactions by low-code/no-code AI platforms' enabling technologies, in which the attribution of moral responsibility is made difficult or impossible due to a combination of technical, legal, social, and moral challenges. The aim of this paper is twofold. Firstly, I investigate the extent to which low-code/no-code AI platforms challenge the current concept of moral responsibility and whether such a concept needs to be adapted to the current socio-technical context. Secondly, I employ an interdisciplinary approach to conceptual engineering to advance a proposal for a new concept of responsibility.

## **Room 2**

### **Medicine and Operationalizing Ethics**

**Chair:** Elisabeth Hildt

#### **An Examination of Doctors' Attitudes Toward Medical AI: A Turkish Example**

Seda Gökçe Turan, *Bahçeşehir University, Turkey, Visiting Professor  
Bournemouth University, United Kingdom*

Can medical ethics be made computable? Machine intelligence already helps health-care staff with a number of tasks. Ethical decision-making, however, has not been handed over to machines. We developed the world's first algorithm for advising on a range of moral dilemma situations that occur in medical institutions. In this talk, we explain how we chose an appropriate ethical framework, why we used fuzzy cognitive maps to set up the advisory system, and how we utilised machine learning to train it. We report on the difficult task of operationalising the principles of beneficence, non-maleficence and patient autonomy, and describe how we selected suitable input parameters that we extracted from a training dataset of clinical cases. The first performance results are promising, but an algorithmic approach to ethics also comes with several weaknesses and limitations. Should one really entrust the sensitive domain of clinical ethics to machine intelligence?

#### **A Feminist View of Medical AI Harm**

Clair Baleshta, *Western University, Canada*

Despite widespread concerns about the harms of Artificial Intelligence (AI) technologies in contexts like medicine, there has been little explicit analysis of what constitutes a 'harm' when looking at the impact of these AI systems (Altman et al. 2018; Smuha 2021). Such an analysis is necessary, however, for both the accurate identification and the effective reduction of harms imposed by this technology. This is a larger task than it may appear to be. The difficulty of providing a clear guiding concept of harm can be seen through an appeal to the philosophical literature – despite harm playing a central role in moral, legal and political philosophy, there is still significant disagreement regarding how to define the concept (Shiffrin 2012). Nevertheless, this paper aims to develop a distinctive, feminist account of AI harm, connecting such harm to a constraint-based view rooted in relational theories of autonomy. Given the primacy of autonomy concerns in medicine, the account will be particularly valuable for understanding the harms of AI in the medical context, as well as clarifying the existing literature surrounding this harm. However, the account will also be applicable to, and have implications for, our understanding of AI harm more generally.

**Building an Algorithmic Advisory System for Moral Decision-Making in the Health-Care Sector**  
<https://journals.library.iit.edu/index.php/CEPE2023/article/view/251>

Lukas J. Meier, *University of Cambridge, United Kingdom*

Alice Hein, *Technical University of Munich, Germany*

Can medical ethics be made computable? Machine intelligence already helps health-care staff with a number of tasks. Ethical decision-making, however, has not been handed over to machines. We developed the world's first algorithm for advising on a range of moral dilemma situations that occur in medical institutions. In this talk, we explain how we chose an appropriate ethical framework, why we used fuzzy cognitive maps to set up the advisory system, and how we utilised machine learning to train it. We report on the difficult task of operationalising the principles of beneficence, non-maleficence and patient autonomy, and describe how we selected suitable input parameters that we extracted from a training dataset of clinical cases. The first performance results are promising, but an algorithmic approach to ethics also comes with several weaknesses and limitations. Should one really entrust the sensitive domain of clinical ethics to machine intelligence?

**11–11:15 a.m.**

**Break**

**CDT 5–5:15 p.m.  
CET**

**11:15 a.m.–1:45 p.m. CDT  
3 5:15–6:45 p.m. CET**

**Session**

**Room 1**

**Relational Aspects of Human-AI Interaction**

**Chair:** Elisabeth Hildt

**Engineering a Concept of AI Neutrality to Protect Against Undue AI Bias**

Roxane Kurtz, *University of Illinois, Springfield, United States*

Algorithmic bias arises when an AI-involved decision results in differential burdens and benefits across groups. Sometimes such bias is morally and politically permissible (or even required). For instance, an algorithm that selects amongst applicants to a mentoring program for members of underrepresented groups should be biased. Thus, we need to distinguish undue algorithmic bias from acceptable bias.

But, both AI Neutrality of Justifications and AI Neutrality of Outcomes fail in this task. Instead, the best candidate for a concept of AI neutrality that properly reveals undue bias is my AI Neutrality of Justified Outcomes. This view requires an entity governing an AI to justify algorithms only by consistent appeal to legitimate goals and to justify differential burdens and benefits due to AI-involved decisions on the grounds that there are no reasonable alternative algorithms for achieving its legitimate goals with less normatively concerning bias.

**The Normative Side of Building Friendship with AI Companions**

<https://journals.library.iit.edu/index.php/CEPE2023/article/view/272>

Tugba Yoldas, *University of Alberta, Canada*

In this paper, I discuss friendship from the perspective of Aristotelian virtue ethics and claim that friendship with AI companions might be harmful to us because firstly, they do not actually meet our moral needs where friendship is concerned, and secondly, they threaten to undermine the virtue of friendship by introducing a morally defective mode of friendship and by transforming important societal norms about interpersonal relationships. I argue that designing ethically-aligned AI companions gains significant importance, especially in the areas where these companions come in contact with vulnerable groups in society such as children, the elderly, and people living with disabilities. To conclude, I suggest that AI companions could prove helpful, for example, as models to help children practice ethical and intellectual virtues, as mental health aids, and as ethical mediators to help us make better moral decisions and navigate ethical dilemmas. But this is a possibility only if their design respects certain ethical principles such as trustworthiness, privacy, safety, and reliability.



### **Experiencing AI and the Relational ‘Turn’ in AI Ethics**

<https://journals.library.iit.edu/index.php/CEPE2023/article/view/237>

Jason Branford, *University of Hamburg, Germany*

The still burgeoning field of AI Ethics, while unprecedented in terms of publication and interdisciplinary, has already come under fire for being sorely limited in important respects. What I here term the relational ‘turn’ in AI ethics seeks to rectify some of these misgivings by emphasizing the rich social context and deeply relational character of human lives (and by extension the injustices that presently plague them and are purportedly exacerbated by AI technologies) as a necessary aspect of any ethical consideration of the use of AI technologies. I endorse yet seek to both bolster and embolden this move. Specifically, I argue that there is a need to (1) deepen the social account of the self that underscores the relational turn to explicitly encompass the “co-constitutive” nature of human-technology relations (specifically, our lived experience of AI technologies) and (2) detail the features and direction of a mature relational ethics (i.e., what it should seek to do and how it ought to go about doing so). This can be achieved, I propose, by turning to and unifying attempts to develop a postphenomenological ethics of technology on the one hand, with the method and intention of a Deweyan moral inquiry (1938) and the pragmatic ethics this has inspired on the other.

### **Social Media Algorithms and Social Distrust**

<https://journals.library.iit.edu/index.php/CEPE2023/article/view/267>

Heather Stewart, *Oklahoma State University*

This paper examines the connection between algorithmic processes on social media and their effects on social life both on- and offline. Specifically, I argue that two algorithmic processes, which my colleagues and I call algorithmic targeting and algorithmic sorting, create epistemic and social conditions that worsen the problems of epistemic injustice and social distrust (see Stewart, Cichocki, McLeod 2023). I aim to show that a robust accounting of the injustice and oppression that algorithms cause or contribute to must also include analysis of the seemingly subtle ways in which algorithms influence our social worlds and daily interpersonal interactions with others, on- and offline. These seemingly “micro” injustices caused or exacerbated by social media algorithms ought to be part of growing dialogues about algorithmic bias and algorithmic oppression.

## **Room 2**

### **Ethics of Data Accumulation and Large Language Models**

**Chair:** Kelly Laas

### **Intercultural Information Ethics Applied to the Data Colonialism Concept**

<https://journals.library.iit.edu/index.php/CEPE2023/article/view/229>

Jonas Ferrigolo Melo, *University of Porto, Portugal*

This paper explores how and to what extent data colonialism can be articulated with Intercultural Information Ethics from exploratory-descriptive research supported by bibliographic research. We connect this approach with the proposal of neo-colonialism, surveillance capitalism, and decolonialism. We conclude that there is a universalized and socially accepted vision of order and wholeness facing the treatment of personal data, categorizing the subjects, and configuring the critique in a patriarchal European idea that sustains the dynamics of modernity. We propose to think about data colonialism from an alternative viewpoint while still pointing out the resistance to data colonialism.

### **Some Preliminary Numbers of Teamwork (Human) Moderation**

Zihe Luo, *Independent Researcher, United States*

Moderating toxic or hate speech online is difficult, but it is very necessary, and, at least in the near future, it requires human intelligence. To improve human moderation, I propose teamwork. In this computer simulation, a moderation hub performs better if all moderation decisions were made by small teams instead of individual moderators. The model also explores what could happen if moderators were given feedback by other moderators. The model suggests that teamwork is expensive but could produce positive effects on reliability. However, teamwork can hurt reliability as well. We will need to find the right way to use teamwork.

### **The Right Response, Not Just the Most Likely: The Limits of Minimizing Error Prediction for Training Large Language Models**

Jacob Browning, *New York University, United States*

The surge in interest in natural language processing has led to a plethora of new, impressive language models. But the outputs of these systems are often inappropriate: dishonest, inconsistent, offensive, and so on. While many attempts have been made to correct these problems, these solutions tend to focus on trying to minimize the statistical likelihood of inappropriate content. But they do little to ensure the system grasps that the content is inappropriate and should not be said—and, thus, the results have been inconsistent. In this paper, I argue these problems stem from these machines regarding only statistically unlikely outputs as inappropriate. Human language users, by contrast, acknowledge many other types of inappropriateness—dishonest, offensive, irrelevant, and so on—which these machines are not being trained to detect. Until these machines become sensitive to these other norms, they will inevitably provide outputs.

### **Registering AI-Generated Patents: A Revolution in Distributive Justice?**

<https://journals.library.iit.edu/index.php/CEPE2023/article/view/277>

Naama Daniel, *Hebrew University of Jerusalem, Israel*

This paper argues that national patent registration policies regarding AI-generated inventions can lead to a new distribution of justice. Currently, some countries register AI-generated inventions as patents while others do not. This difference may affect the distributive justice problem in patents, specifically regarding medicine, which is identified by a twofold inequality/ inequity towards developing countries: first, R&D is undertaken to benefit rich developed countries, and not to cure illnesses common in developing countries; second, the monopolistic price of medicine render it inaccessible for purchase by developing countries. By deciding whether or not to register AI-generated inventions as patents, developing countries now have the unique opportunity of influencing and even changing this inequality/ inequity. They can register such inventions, thereby incentivizing AI-operators to direct their AI at inventing cures to illnesses common in their territories, or refuse to register them, thereby avoiding the need to pay a monopolistic price for any AI-generated invention.

## Day 2—May 25

8:00 a.m. CET/4 p.m. KST  
6 p.m. AEST/8 p.m. NZST

### **Session 4**

#### **Room 1**

### **Ascribing Human Characteristics to AI: Advantages and Pitfalls**

**Chair:** Elisabeth Hildt

#### **Smart Machines and Wise Guys: Can Intelligent Machines be Wise?**

Edward Spence, *University of Sydney, Australia*

This paper will address an important topic in the Philosophy and Ethics of Technology, which is that of artificial mediated agency and autonomy and its moral significance and impact on human wellbeing. Furthermore, it will address the key question of this paper whether artificial agency and autonomy can be extended to the notion of wisdom: if agency is mediated and distributed between humans and intelligent machines, can there also be a mediated wisdom?

#### **Interpreting Ordinary Uses of Psychological and Moral Terms in the AI Domain**

<https://journals.library.iit.edu/index.php/CEPE2023/article/view/255>

Hyungrae Noh, *Sunchon National University, South Korea*

Ordinary speakers commonly use emotion-terms or moral-patency-terms to explain the behavior of AI robots. This paper is about whether this referential shift from the human domain to the AI domain entails semantic changes: do ordinary speakers literally consider AI robots to be emotional beings or moral patients? Two non-literalist accounts for semantic changes concerning these terms used in the AI domain will be discussed: the technical view (ordinary speakers express technical senses) and the emotion view (ordinary speakers express their own affective empathetic states). I will examine whether these non-literalist accounts are supported by relevant empirical experiments, and I argue

that the emotion view offers a plausible interpretation. This argument suggests that the judgments made by ordinary speakers regarding the appropriate domain of these terms are fallacious as ordinary speakers fail to express literal meanings of the terms in the AI domain.

### **Can We be Friends with AI? What Risks Would Arise from the Proliferation of Such Friendships?**

<https://journals.library.iit.edu/index.php/CEPE2023/article/view/254>

Nick Munn and Dan Weijers, *University of Waikato, New Zealand*

In this paper we analyse friendships between humans and artificial intelligences, exploring the various arguments that have been or could be offered against the value of such friendships, and arguing that these objections do not stand up to critical scrutiny. As such, we argue that there is no good in-principle reason to oppose the development of human-AI friendships (although there may be some practical defeasible reasons to worry about such friendships becoming widespread). If we are right, there are important implications for how friendship is conceptualised and valued in modern times. Furthermore, if human-AI friendships are in-principle valuable, the moral responsibilities for how governments and corporations should act in regards to AI friends are quite different to those generated by human-AI friendships being dis-valuable.

### **Moral Attribution in Moral Turing Test In-Person**

Jolly Thomas and Mubarak Hussain, *Indian Institute of Technology Dharwad, India*

To assess the intelligence of machines, Alan Turing proposes a test of imitation known as the imitation game and famously known as Turing Test. To assess whether Artificial Intelligent systems could be moral or not, Colin Allen et al. developed a test of imitation in the context of morality, a test known as the Moral Turing Test (MTT). Arnold et al. argue against any type of MTT (Arnold et al. 2016. *Ethics and Information Technology* 18, 103-115). Allen et al.'s attempt to develop the MTT on the one side and the criticisms raised by Arnold et al. on any type of MTT on the other side point towards the distinction between the very attribution of morality and the performance of the morally significant actions to which morality is attributed. Whether to develop an MTT or argue against any type of MTT, such distinction as a component into the analysis must be encompassed.

6:45 a.m. CDT  
12:45 p.m. CET

## **Break**

7 a.m.–9 a.m.  
CDT 1–3 p.m. CET

## **Session 5**

### **Room 1**

### **Ethics and Technological Transformation**

**Chair:** Jason Branford

### **Acceleration AI Ethics and the Dispute between Stability AI's Diffusion and OpenAI's Dall-E**

James Brusseau, *Pace University, United States*

One objection to conventional AI ethics is that it slows innovation. This presentation responds by reconfiguring ethics as an innovation accelerator. The critical elements develop from a contrast between Stability AI's Diffusion and OpenAI's Dall-E. By analyzing the divergent values underlying their opposed strategies for development and deployment, five conceptions are identified as common to acceleration ethics. Uncertainty is understood as positive and encouraging, rather than discouraging. Innovation is conceived as intrinsically valuable, instead of worthwhile only as mediated by social effects. AI problems are solved by more AI, not less. Permissions and restrictions governing AI emerge from a decentralized process, instead of a unified authority. The work of ethics is embedded in AI development and application, instead of functioning from outside. Together, these attitudes and practices remake ethics as provoking rather than restraining artificial intelligence.

### **Swarm Ethics: A New Collective and Decentralized Ethics in the Digital Age**

Katja Rausch, *House of Ethics, Luxembourg*

Daniele Proverbio, *University of Luxembourg*

Swarm Ethics is a new multi-disciplinary approach to ethics for the "individual-within-open-societies", based on principles of swarm intelligence, open systems and digital social technologies.

It transcends traditional cognitive-based ethical paradigms and considers perception-action behavior by humans as the driving force in fast-paced, modern ethical decision-making. Swarm Ethics identifies multitudes of decentralized intelligent agents that manage to perform coordinated, self-organized ethical activities by sharing information and acting upon dynamical environments. Swarm Ethics is thus “emerging” collective ethics at large scale and fast pace, fueled by digital technologies that speed up all processes.

Swarm Ethics fully acknowledges humans as dynamic and responsive agents in open, mutant systems. Mutuality, imitation, empathy and care are paramount for Swarm Ethics to happen. The emerging ethical behavior is purpose-driven without being “hard-coded” in agents. The novel concept can be a game-changer when apprehending the rise of collaborative and companion robots, interactions in web3 and metaverses or when analyzing digital polarization and political/ideological phenomena. Within this novel paradigm, ethics is fit for the digital age. It is fully recognized as “difficult” but, at the same time, it regains its prestige: if everything influences ethics, ethics influences everything.

### **Epistemology and Algo-reliabilism: A Pathway to Sound Ethical Artificial Intelligence**

<https://journals.library.iit.edu/index.php/CEPE2023/article/view/257>

Helen Titilola Olojede, *National Open University of Nigeria*

Artificial intelligence (AI) is a topical issue in the world and in academic discourse. As the clamour for ethical regulations increase, certain questions that should be asked include: Is ethics or ethical principle for the originators of AI or for AI itself or for the users of AI? What role does epistemology play in the ethical discussion of AI? How should we understand individual ethics questions? How about collective or societal ethics question? This paper thus, clarifies some of these questions, and based on the importance algorithm plays in the life of AI system, argues for what it calls algorithmic reliabilism (AR) in the development of AI. AR conceived as a set of normative accountability mechanisms, incorporates Goods such as solidarity, human dignity, interpretive fineness, coherence to make AI safer and more just while becoming more efficient for persons as individuals and groups of persons.

### **On the Possibility of Moral Machines: A Reply to Robert Sparrow**

<https://journals.library.iit.edu/index.php/CEPE2023/article/view/243>

Dane Leigh Gogoshin, *University of Helsinki, Finland*

Robert Sparrow has recently argued (Sparrow 2020) against the possibility of building moral machines (or AIs). He urges the community to better understand the nature of ethics before undertaking this endeavor. He suggests, following philosopher Raimond Gaita, that ethics is much more complex, more personal, and in many ways more subjective than is generally acknowledged in machine/AI ethics, and he deploys original thought experiments to show this. In this paper, while granting Sparrow’s overall contention – that ethics is too complex to presuppose the possibility of moral machines – by way of challenging certain of Sparrow’s premises and triggering different intuitions from his thought experiments, I defend a cautiously optimistic view of the possibility of moral machines/AIs.

### **Room 2:**

### **Conceptual Reflections and AI**

**Chair:** Kelly Laas

### **Ex-Post Approaches to Privacy: Trust Norms to Realize the Social Dimension of Privacy**

<https://journals.library.iit.edu/index.php/CEPE2023/article/view/287>

Haleh Asgarinia, *University of Twente, Netherlands*

Privacy is a social value because it cannot be constituted individually but rather depends on relationships that involve many, and it is not individually valuable, even though it is enjoyed by individuals, because its value derives from interaction with those who share interests in caring for what one cares about. It follows that what matters is not who has access to personal information or if a person is capable of exercising control over their personal data, but rather who cares about shared data, how to care about it, and to whom a person cares to reveal their personal data. Given that trust includes care and, trust is the constituent of the social value of privacy, this paper will address this question: “what are trust norms to be trustworthy?” This paper aims to identify trust norms in the context in which personal data are shared, processed by an AI system, and



information is inferred from the shared data. I consider the following argument to identify norms to be trustworthy in response to the other's trusting reliance. The first premise is that a person who is motivated to act does not make trust appropriate. Second, it is argued that the condition that gives rise to trustworthiness resides in the stance the trustor takes toward the trustee. Third, it is argued that norms for the trustee to be trustworthy derive from norms on promise-making. And finally, it is argued that trusting others' words involves relying upon them to fulfil promise-making norms. I conclude that a person who is responsible for answering a question while using an AI system would be trustworthy only if first, s/he intends to say the correct answer; second, s/he is competent to answer the question; third, s/he, in fact, asserts the right answer, and finally, the AI system is accurate or epistemically reliable. In this regard, promoting trust norms of sincerity and competence by the trustee and developing an accurate AI system by designers both contribute to constructing privacy and realizing its social value.

### **Why is an Intervention Epistemically Paternalistic?**

Lei Niu, *University of Cologne, Germany*

There are widespread misinformation and disinformation. As a result, many people tend to question their true beliefs and hold false beliefs. On the one hand, it is always ineffective to delete all harmful information, because a non-perfect definition can exclude some true information and legitimate some harmful information. On the other hand, it is not expected that information consumers can timely improve their epistemic abilities, since the formation of true beliefs and knowledge requires expertise that is burdensome. Countermeasures, including AI technology-based epistemic interventions, against those consequential epistemic phenomena are necessary.

The wide applications of information technology that aim at countering disinformation and fostering true beliefs have an epistemic paternalistic profile since formal consent is always absent, especially when new information technology is applied to groups that include more than one person. This enriches and complicates the discussion of paternalism. Epistemic paternalism faces two challenges: a conceptual challenge and a normative challenge. First, one may worry that epistemic paternalism is not distinctively epistemic, and it is just one kind of general paternalism. Second, one worry is that a distinct epistemic paternalism cannot be justified. This paper will clarify these two challenges in detail and attempt to respond to them.

### **Understanding Artificial Intelligence, and Virtue Ethics**

<https://journals.library.iit.edu/index.php/CEPE2023/article/view/249>

Mahdi Khalili, *Institute for Research in the Fundamental Sciences, Iran*

In artificially intelligent systems that are unintelligible, it is unclear why specific patterns have been extracted from a given dataset. This paper argues that unintelligible technologies, including unintelligible artificial intelligence, have the potential to require that human users be indifferent to the understanding of reasons for decisions and actions. This entails that unintelligible artificial intelligence has the potential to require that its users do not realize their moral capacity. Then, the paper draws on the literature on scientific understanding to suggest that an artificially intelligent system can be rendered understandable if a qualitative account of the consequences of its use in context is provided.

**9–9:15 a.m. CDT**  
**3–3:15 p.m. CET**

**Break**

[Room 1](#)

**Healthcare and Explicability**

**Chair:** Elisabeth Hildt

**AI Explicability in Medicine and Healthcare: Fighting Against the Return to the Paternalism**

<https://journals.library.iit.edu/index.php/CEPE2023/article/view/253>

Lorella Meola, *University of Salerno, Italy*

The use of Artificial Intelligence (AI) in medicine and healthcare increases the efficiency and accuracy of clinicians and their medical decision-making process; at the same time, it raises several ethical challenges too. In particular, we are going to analyse the issue of explicability of AI system in an ethical perspective.

Notably, clinicians can be assisted or even replaced by AI systems in the decision making process, improving care quality, that means care accuracy, and optimizing the time spent with the patients.

However, the lack of transparency and the opacity of AI could mine the therapeutic relationship; thus, it could weaken trust, discourage patients and destroy the basis of their self-determination.

We are going to outline that a lack of algorithmic transparency could compromise patient autonomy, a pitting point of biomedical ethics since the second post-war period. If we are not entitled to ask for explanations about AI decisions, we are going to interact with a system that claims to know what is good for us because it knows health standards, introducing a new form of paternalism, that could have ethical effects on the patients and society too.

**Trust through Explanation? On the Claim for Explainable Medical Decision Support Systems**

<https://journals.library.iit.edu/index.php/CEPE2023/article/view/263>

Sebastian Scheidgen, *FernUniversität in Hagen, Germany*

Artificial intelligence (AI)-based systems will play an increasingly significant role in clinical practice. For example, decision support systems (for tumor identification and classification in diagnostic imaging) or AI-based systems for automatic data analysis with regard to diagnosis, prognosis and prediction are already being developed and increasingly used.

On the one hand, this is associated with the hope that AI will be able to answer certain questions faster, more efficiently, and more effectively than human physicians. On the other hand, a number of ethically relevant problems is raised by the eventual use of such technologies, in particular issues of quality, safety and unfairness, which are primarily reflected in systematically inaccurate or erroneous results of AI-based systems. In addition, the risk for so-called “workflow disruptions” has been documented for contexts of using AI-based decision support systems.

It has been argued that these problems may have a negative impact on the doctor-patient relationship or even patient trust in medical care. A frequent proposal to address these issues consists in the demand for explainable deep learning-based decision support systems, i.e. systems whose operations, roughly speaking, can be understood by humans. In the proposed talk, I will address the question as to whether (or in what sense) this proposal may succeed, or as to whether (or in what sense) explainable medical decision support systems may contribute to a good doctor-patient relationship.

**AI Opacity vs. Patient’s Autonomy in Decision-making**

<https://journals.library.iit.edu/index.php/CEPE2023/article/view/247>

Jose Luis Guerrero Quiñones, *Oxford Brookes University Spain*

The use of Artificial Intelligence (AI) in healthcare contexts is highly controversial for the (bio)ethical conundrums it creates. One of the main problems arising from its implementation is the lack of transparency of Machine Learning (ML) algorithms, which are thought to impede the patient’s autonomous choice regarding her medical decisions. If the patient is unable to clearly understand why and how an AI algorithm reached certain medical decision, her autonomy is being hovered. However, there are alternatives to prevent the negative impact of AI opacity in shared (healthcare professional - patient) decision-making processes, and benefit from the high accuracy of such systems. Despite the ethical issues arising from the lack of transparency of AI systems and their influence in patients’ decision-making, the paper advocates for its usage and defends the positive

role that it could have in enhancing patients' autonomy. We advocate for a system where the implementation of AI is only partial, and humans will remain in control of overseeing algorithm interpretation of images and patients' data.

### **Transparency and Authority Concerns about the Use of Algorithmic Ethical Decision-Making in Healthcare**

Michael Robinson, *Chapman University, United States*

Jeffrey Byrnes, *Grand Valley State University, United States*

In response to recent proposals to utilize machine learning to automate ethics consultations in health care, we raise two problems—the transparency problem and the authority problem—for the prospect of having medical professionals rely on algorithms to provide ethical guidance in clinical matters. The first cause for concern is that, because these algorithmic recommendations would effectively function like black boxes, this approach seems to preclude the kind of transparency that would allow medical staff to explain and justify treatment decisions to patients, fellow practitioners, and those tasked with providing oversight of those recommendations. The other problem is that the kind of authority that would need to be given to the guidance issuing from these programs in order to do the work set out for them would mean that medical staff will not be empowered to set aside this guidance or provide any meaningful check against it in those cases when its recommendations are morally problematic. Taken together, these concerns provide sufficient reason to think that algorithms will not be suitable for replacing human beings in making ethics recommendations in health care.

## **Room 2**

### **Intelligent Agents and Designing the User Experience**

**Chair:** Ray Trystad

#### **MIIND and HEART: Measuring and Designing for Thicker Qualities of User Experience**

Rachel Siow Robertson, *Hong Kong Baptist University, Hong Kong*

Jennifer George, *Goldsmiths University of London, United Kingdom*

Matthew Kuan Johnson, *University of Oxford, United Kingdom*

User experience of digital platforms and technologies tends to be quite 'thin', characterized by low-quality engagements such as addictive tendencies or browsing on autopilot. This paper takes an interdisciplinary approach to identifying the cognitive and active dimensions of 'thick' user experience, introducing a new design and UX framework which is centered around the notion of joy. We first discuss existing frameworks for measuring the quality of user experience. We explain how many harmful kinds of engagement either do not impact negatively on usual metrics, or even result in higher scores according to these assessments. We then draw on recent work in positive psychology, which has seen a move from happiness and 'thin' conceptions of pleasure to 'thicker' notions such as joy. We use these resources to develop the MIIND framework as a supplement to usual approaches, with the dimensions of Motivation, Integrity, Intensity, Normative Content, and Dependent. We also provide examples of survey questions for measuring and evaluating these dimensions. We end by suggesting new research directions, including exploring how MIIND metrics correlate with increases in adoption, retention, task success, and financial profitability.

#### **Do Children Dream of Connected Watches? How the Connected Citizens Experience the World**

<https://journals.library.iit.edu/index.php/CEPE2023/article/view/261>

Lisa Roux, *Université de Pau et des Pays de l'Adour, France*

As our society relies increasingly on artificial intelligence in day-to-day life, we have very limited knowledge and control of its uses and consequences on our representations, values, behaviors, lifestyles, etc. In particular, connected objects take on many different tasks and are deemed to serve us, but they also deeply affect and shape our relationship to the world (e.g. how we interact with others and our environment, what control we have on our sleep) and influences our behaviors. In what ways? What consequences do they have on the way we take decisions and these decisions? In this presentation, I study the power relations between humans and connected objects and identify some ethical issues they raise.

### **Stop at Red? Engineering Meets Ethics**

<https://journals.library.iit.edu/index.php/CEPE2023/article/view/281>

Ignacio D. Lopez-Miguel, *TU Wien, Austria*

Over the past few years, artificial intelligence has fueled a revolution in several scientific fields. Intelligent agents can now give medical advice, translate spoken language, recommend news, and drive different types of vehicles, to name but a few. Some of these agents need to interact with humans and, hence, need to adhere to their social norms. Safety engineers have always worked with critical systems in which catastrophic failures can occur. They need to make moral decisions in order to keep the system under some acceptable risk level. In this paper, we will propose an approach to give a value to contrary-to-duty behaviors by introducing a risk aversion factor. We will make use of decision theory with uncertain consequences together with a risk matrix used by safety engineers. We will successfully exemplify this approach with the problem in which an autonomous car needs to decide whether to run a red light or not.

### **Robots, Wrasse, and the Evolution of Reciprocity**

Michael Dale, *Eindhoven University of Technology, Netherlands*

Due to its prominent role in human sociality, robotics researchers have increasingly considered to what extent reciprocity might be important in human-robot interaction, and whether it should be included as a design feature in social robots. However, very little has been said of the original function of reciprocity. Indeed, as many evolutionary biologists know, reciprocity evolved to foster cooperation among human groups, yet this fact has for the most part remained unexplored in the robotics literature. In this paper, I aim to change that. Specifically, I examine what we know about the evolution of reciprocity in humans and consider to what extent this knowledge can weigh in on discussions about social robots. I argue that the evolutionary account of reciprocity not only helps deflect arguments against using reciprocity as a design feature in social robots; it also reveals that reciprocity is a more complex phenomenon than is currently represented in the robotics literature. Indeed, there is direct reciprocity and indirect reciprocity, and robot designers should take both into account if they want to most effectively enhance human-robot relations.

**11:15-1:30 a.m. CDT**  
**5:15-5:30 p.m. CET**

### **Break**

**11:30-1:30 p.m. CDT**  
**5:30-7:30 p.m. CET**

### **Session 7**

#### **Room 1**

#### **AI Governance and Democracy**

**Chair:** Kelly Laas

#### **Digital Transformations of Democracy: Requirements for Successful Problem-Solving in the Age of Anthropocene**

**URL:** <https://journals.library.iit.edu/index.php/CEPE2023/article/view/275>

Jan-Philipp Kruse, *University of Hamburg, Germany*

Digital transformations of society in general, and of its democratic regulation in particular, are being intensively researched. With “digital constellation,” (Berg, Rakowski & Thiel 2020) a concept has been introduced that discusses decisive factors for the future shaping of democracy.

With my approach, I propose going one conceptual step further from there. A socio-philosophical view of digitalizing democracy can focus even more on the consequences of digital transformations for the functioning of democratic self-regulation. In other words, it is about which development paths make a successful further development of democratic practices likely, and which do not.

On the one hand, this is about drawing a connection between the infrastructure of public spheres, and the quality of democratic discourse.

On the other hand, the structure of problems that are debated in public spheres has also changed. Now, moves to treat problems, or not to treat them, like with climate change, can have irreversible consequences. It can therefore be said that the bar for democratic understanding and collective action is now much higher.

Against this background, the question of development paths is posed differently: it is not just



about a somehow functioning democracy, but a discourse that successfully addresses ecological challenges.

### **Democratic Culture and the Automation of Information: What is Really at Stake?**

<https://journals.library.iit.edu/index.php/CEPE2023/article/view/236>

Jason Branford, *University of Hamburg, Germany*

Eloise Soulier, *University of Hamburg (Germany)*

Laura Fichtner, *University of Hamburg (Germany)*

In recent years, the threats that the algorithmic ordering and manipulation of information in the digital sphere purportedly pose to democracy have received considerable academic attention. In our paper, we add to this work and investigate how the automation of information online—exemplified by the recent advancement and uptake of generative-AI technologies—impact our “democratic culture”; a notion we develop based on a Deweyan account of democracy. This notion goes beyond institutionalised forms of democratic governance and builds on the idea that for democracies to flourish, certain ‘spaces’ for, and practices and norms of, social interaction need to obtain and be nurtured. These enable, for example mutual recognition, the chance for creativity, and equal possibilities for meaningfully participation in societal decision-making processes. Each of these entails both individual relationships to relevant information as well as information-based interrelationships between people who, in recognising and facing collective ‘problems’, constitute a democratic community. Based on this notion, we explore the questions of how the adoption of generative AIs such as ChatGPT might impact the ability for a democratic culture to develop and flourish and, further, what it would mean to design and govern such technologies in a way that supports rather than endangers the proliferation of such a democratic culture.

### **Power with(out) Responsibility? A Comparative Analysis of ‘Effective Accountability’ Across Public Algorithmic Governance**

Ido Sivan-Sevilla, *University of Maryland, United States*

Aileen Nilsen, *ETH, Switzerland*

David Eliot, *University of Ottawa, Canada*

How and why do public agencies limit their use of algorithmic tools in response to accountability demands? Which algorithmic accountability efforts across political, legal, and social forums are likely to succeed? Why? This project investigates successes and failures in altering the deployment of algorithmic tools by public agencies. Understanding accountability as a relationship between public agencies and various ‘accountability forums,’ this project studies accountability interactions during the post-deployment phases of public algorithmic tools for governance. We first develop ‘Algorithmic Accountability Metrics,’ and conduct ‘high-level’ classification of algorithmic accountability cases across national and institutional settings. We detail (1) advocacy strategy; (2) government response to accountability demands; (3) issue saliency of accountability demands; and (4) observed accountability outcomes. We gauge on the variability of ‘accountability outcomes,’ (our dependent variable), which can range from restrictions on the use of the data / the algorithm, temporary or permanent bans on the data or the algorithm, and varying levels of access to the data and the algorithm. Algorithmic accountability is found to be nuanced and multi-dimensional, with varying advocacy strategies, government responses, and issue saliency correlated to different ‘accountability outcomes.’

## **Room 2**

### **AI, Surveillance and Moral Growth**

**Chair:** Tugba Yoldas

#### **Surveillance Culture and Fundamental Rights: The Excluded and the Beneficiaries**

<https://journals.library.iit.edu/index.php/CEPE2023/article/view/229>

Camila Costa, *Federal University of Rio de Janeiro, Brazil*

Jonas Ferrigolo Melo, *University of Porto, Portugal*

Surveillance has progressively grown in social life in the 20th and 21st centuries. It happened partly because of the adoption of multiple sensors that can extract, collect and analyze an enormous volume of data. This expressive data volume, variety, and processing velocity are known as big data. The increasing adoption of big data and models based on algorithmic intelligence has a massive impact on society because of its dissemination among social spheres through relations between the public and private sectors. This paper aims to discuss surveillance culture and its consequences on fundamental rights such as privacy and freedom of speech. In addition, it is intended to debate the excluded and the beneficiaries of a surveillance society. The methodological approach is the literature review. The conclusion relies on the need for intercultural ethics to strengthen the right to privacy to guarantee not only itself but multiple fundamental rights nowadays.

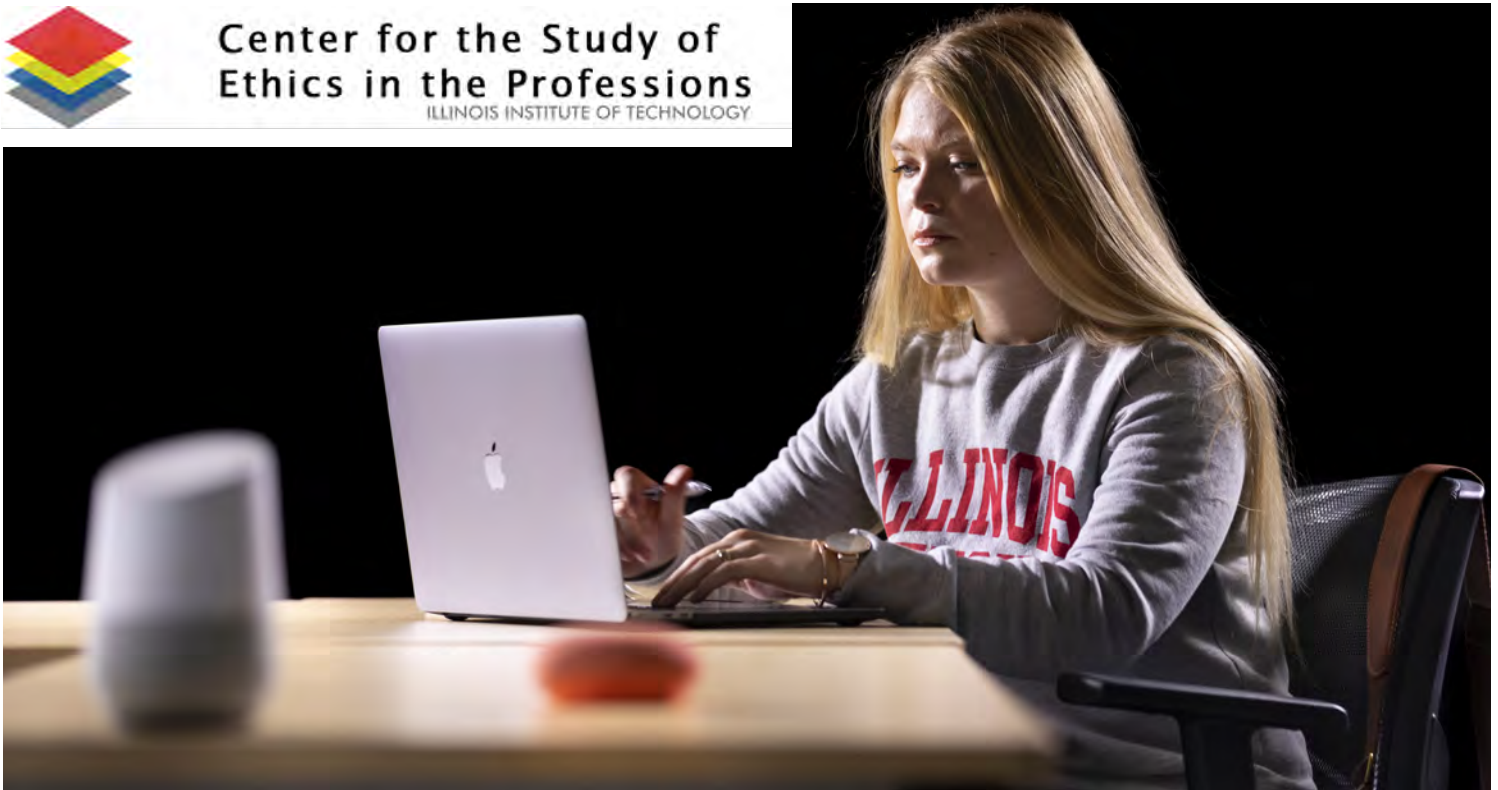
#### **Artificial Intelligence and Moral Growth**

Adam Zweber, *Stanford University, United States*

Taking a cue from bioethics, the contemporary literature on the ethics of artificial intelligence has seen burgeoning discussion of the potential to use AI to enhance human moral decision making. In this paper, I argue that an underappreciated danger of the use of such technology is that it would inhibit the user's moral growth. In particular, it would discourage what John Stuart Mill calls "experiments of living." Users would likely fail to undergo "transformative experiences" that would lead them to reflect on and revise their values. They would also be less likely to engage in and learn from Millian experiments of living. The outcomes of our "experiments" would even be effected by the fact that an AI recommended a certain course of action to us. This, I claim, would greatly reduce the value of such experiments, and is telling against several prominent proposals for the use of AI to morally enhance humans.

**1:30-1:45 p.m. CDT**  
**7:30-7:45 p.m. CET**

**Closing**



## MISSION:

To educate students as responsible professionals, to reflect on the wider implications of scientific progress, and to contribute to the shaping of technology in accordance with fundamental human values.

## VISION:

CSEP will be an internationally connected ethics center with a focus on professional and applied ethics, integrating ethics education into the colleges and departments of Illinois Tech, engaging in research and public dialogue at a local and global level.

## WE SEEK TO:

- Enhance the distinctive education offered to Illinois Tech students by working with faculty from all the different colleges and departments at Illinois Tech to help meaningfully integrate ethics into their educational programs – from the undergraduate to the graduate level.
- Promote innovative teaching by developing new pedagogical approaches and content in a wide variety of formats from the semester-long ethics course to shorter lessons, workshops or other formats. Establish a strong research program in ethics in the life sciences and in ethical and societal issues of emerging technologies.
- Build on the already existing unique CSEP collection of codes of ethics, expand and internationalize the collection to make it the basis for future research on codes of ethics.
- Be a strong participant in debates on ethical and societal implications of science and technology in the Chicago area, nationwide and internationally.

The Center for the Study of Ethics in the Profession's research program focuses on ethics in the life sciences and ethical and societal issues of emerging technologies, with a particular focus on philosophical and ethical aspects of neuroscience. The Ethics Center is committed to multi-disciplinary and multi-institutional research, to projects that combine empirical investigation with conceptual analysis, and to projects that introduce and propagate innovations in teaching. Furthermore, the Ethics Center Library houses a unique collection of ethics codes from all over the world and a large collection of ethics education materials.

Externally funded projects enable CSEP to conduct interdisciplinary research involving practitioners, as well as academics from Illinois Tech and other institutions. Topics CSEP has addressed include ways in which the brain and behavioral sciences might provide insight into moral and philosophical questions, intellectual property protection for science and technology, national security restrictions on the dissemination of scientific and technical information, responsible research and innovation in science, university/industry research relationships, organizational development, ethics in vocational education, and individual and collective responsibility in engineering.