KEYNOTE SPEAKERS

Patrick Jagoda, Dept. of English & Dept. of Cinema & Media Studies, University of Chicago, USA**:**

How Video Games Can Help Us Think Through Networks

Artificial intelligence research has increasingly moved computation from matters of symbolic manipulation and calculation, and toward heuristic methods of testing that enable neural networks to adapt through the very process of learning. Transformations in both digital and networked technologies, including in the realm of AI, have also enabled substantive changes in culture. Since the late twentieth century, one of the most quickly developing cultural forms has been the video game. As of 2016, digital games engaged an estimated 1.6 billion people worldwide. Though games are usually considered a popular medium, they can be characterized as "experimental" in a number of senses that resonate with both avant-garde art and scientific hypothesis testing. Through their interactive and participatory affordances, games operate as models of systems and networks that enable real-time experiments with human behavior and decision-making. This talk explores video games as an important experimental form at a historical moment characterized by digital media, deep learning algorithms, and telecommunication networks — as well as the forms of distributed cognition and affective labor that these technologies occasion. In a number of ways, games alter how we approach networks as a way of thinking and acting within twenty-first century culture.

Short bio

Patrick Jagoda is Associate Professor of English and Cinema & Media Studies at the University of Chicago. He is also a co-editor of Critical Inquiry and co-founder of both the Game Changer Chicago Design Lab and the Transmedia Story Lab. He is the author of Network Aesthetics (2016) and co-author with Michael Maizels of The Game Worlds of Jason Rohrer (2016). He is currently working on his next book, Experimental Games. Patrick has published over thirty essays in both humanistic and scientific journals. For more on this projects and writing, please visit: http://patrickjagoda.com/.

Maria Schuld, Quantum Research Group, University of KwaZulu-Natal, Durban, South Africa & Xanadu Quantum Computing Inc, Toronto, Canada

Quantum Neural Networks

Quantum neural networks is a term that summarises approaches to combine neural networks with quantum physics. In the 1990s, these were mostly inspired by biological neural nets and the attempt to explain dynamics of the brain using quantum theory - giving rise to many controversial debates and open questions. Since recent years, quantum neural networks are increasingly known as algorithms for quantum computers that implement variations of artificial neural networks. Quantum computers are machines that process information based on the laws of quantum mechanics, and spark a lot of commercial interest as the first generation of devices is being developed and tested in academic and industry labs around the globe. An entire research field looks at using these devices for machine learning tasks, and neural networks such as Boltzmann machines, Hopfield networks

and multi-layer perceptrons play a prominent role in this relatively new branch of research. This talk will give an overview of both quantum perspectives on biological and artificial neural networks to an interdisciplinary audience, and attempts an outlook of what to expect from quantum physics in future.

Short bio
Maria Schuld did her PhD in physics at the University of KwaZulu-Natal (South Africa) in 2017 as a fellow of the German Academic Foundation. In her dissertation she looked at ways of how to combine quantum computing and machine learning, with a particular focus on artificial neural networks. She previously obtained a MSc degree in physics at Technical University of Berlin, and a Diplom in political science at Freie University of Berlin. She is currently working as a Post-Doc at the University of KwaZulu-Natal and freelances for the Toronto-based quantum computing startup Xanadu as well for the Open Data Durban civic technology lab.

José del R. Millán, Swiss Federal Institute of Technology in Lausanne (EPFL), Switzerland

Brain-Machine Interfaces: A Tale of Two Learners
Real-time signal processing and decoding of brain signals are certainly at the heart of a brain-machine interface (BMI). Yet, this does not suffice for subjects to control a BMI. In this talk I will review some of our recent studies, most involving users with severe motor disabilities, that illustrate additional principles of a reliable BMI. A central theme is the need for a comprehensive mutual learning methodology that reinstates the three learning pillars of a brain-controlled device (at the machine, subject, and application level) as equally significant. I will also show how, in the case of rehabilitation, a BMI does not necessarily need to aim and achieve the highest possible decoding performance to be effective. Finally, I will discuss recent work in our laboratory illustrating how to enhance subject learning and BMI performance through appropriate feedback modalities.

SPEAKERS

Mads Sloth Vinding, Department of Clinical Medicine, Aarhus University, Denmark

Deep learning MRI pulse sequences
Deep learning has entered the field of magnetic resonance imaging (MRI) until now primarily for reconstruction of images from raw data and diagnostic purposes, e.g., tumor grading. Here, we take deep learning under the hood of the MRI system. An MRI experiment consists, simply speaking, of roughly four stages; 1) a preparatory stage, where the magnetization is excited by radio frequency (RF) pulses to a measurable state; 2) a mixing stage, where the magnetization is encoded with an image contrast of choice, e.g., diffusion weighting; 3) a measurement stage, where the magnetization's signal is stored; and 4) a magnetization equilibrium recovery stage. The four stages are cycled until an entire MR image can be reconstructed and analyzed. In this work, we are, with deep learning, pursuing a special kind of 2D RF pulse waveforms, which inheres to the first stage. They excite the magnetization in a 2D pattern of choice, e.g., shaped like a specific brain region. These pulses are normally optimized in a computer, whenever the shape and other characteristics of the pulse become infeasible to solve analytically. Yet, the computer optimization can be strenuous even on modern commodity hardware. We show it is possible to deep learn a neural network for completing the task of producing a 2D RF waveform based on the input of the 2D pattern of choice.

Short bio
Mads Sloth Vinding (MSV) worked on Fast Optimal Control in MRI in his PhD project defended in 2012. Since then MSV has worked on further developing pulse sequences for MRI and NMR,

involving long-lived singlet-state storage for dissolution-DNP by optimal control, and MRI pulse sequences for novel state-of-the-art multi-channel transmit MRI facilities. MSV has experience with 19F MRI of targeting/drug-carrying nano-particles, and single-crystal NMR hardware, from his time at Interdisciplinary Nanoscience Center, Aarhus University. MSV is currently employed as Assistant Professor at the Center of Functionally Integrative Neuroscience at the Department of Clinical Medicine, Aarhus University, where his main project concerns hyperthermia treatment in brain tumors.

Simon Enni, Department of Computer Science, Aarhus University, Denmark

The Case for Contestability in Intelligent Computer Systems

The increasing prevalence of automated decision-making by intelligent and autonomous computer systems raises serious ethical and legal issues. When an intelligent computer system acts automatically, the stakeholders of the system might suspect that its behavior is not aligned with their intensions, e.g. when doctors suspect that an automated medical diagnostics system is making unsound inferences. Alternatively, the people affected by automated decisions might dispute the legitimacy of the system's decision-making authority, e.g. if they suspect that a system is discriminating against a minority group. While human decision-makers can be interrogated to provide reasons and justifications for their decisions, few intelligent computer systems provide similar services, and systems using neural networks in particular are known to be opaque in this sense. A central argument is that in order to justify intelligent computer systems having decision-making authority, we should be able to legitimate this authority. To this end, we can be inspired by how human decision-making authority is legitimated. A requirement of non-opacity (Machin 2009) is often used to rule out political decision-makers who either cannot or will not explain or justify their decisions, tying the legitimacy of decisions to the degree to which they can be explained. Some also insist that legitimate decision-making authority requires procedures in which people can participate and deliberate. I argue that intelligent computer systems can be designed to be contestable by enabling users to engage critically in understanding and challenging the behavior of the system. A contestable design could address the requirement of non-opacity and involve those concerned in a critical investigation of the system, thereby tracing a possible direction for legitimating autonomous decision-making. I will present a conceptual framework for developing contestable systems as well as recent related developments in the field of Explainable AI.

Short bio

Simon Enni is part of the Data Intensive Systems group at the Department of Computer Science at Aarhus University working on the interdisciplinary FounData project. He received his Bachelor of Computer Science from Aarhus University and has since then been enrolled in the PhD programme at the Department of Computer Science. His research interests includes data mining, machine learning, and methods enabling us to better understand and critically engage with the behavior of automatically derived models from these fields.

Michael Castelle, Centre for Interdisciplinary Methodologies, Warwick University, UK

Deep Learning and the Coming Crisis for Social Theory

The contemporary field of *deep learning*, which uses convolutional and/or recurrent neural network structures to recognize and generate content from image, text, speech, and other forms of predominantly unstructured data, has been described by practitioners as a "revolution" (Chollet, 2017) and scientific "tsunami" (Manning, 2015); it is the most recent and perhaps most dramatic resurgence of the subsymbolic paradigm, after its previous rise (Rosenblatt, 1958) and fall (Minsky and Papert 1968) and rise (Rumelhart and McClelland 1986) and fall (Vapnik 1998). In this paper, I will argue that the distinctiveness of deep learning derives from its *ensembling* of epistemologies

and techniques, combining elements of older behaviorist, cognitivist, structuralist, connectionist, and predictive machine learning paradigms alongside layered architectures, objective functions, thoroughly-'vectorized' data sieves (Mackenzie, 2017; Kockelman, 2013), and exteriorized supply chains of both human and machine labor (Scholz, 2012). While the resultant 'conquest of iconicity' at a spatiotemporal multiplicity of perceptual levels has radically transformed the fields of computer vision and machine translation, I argue that the dual qualities of deep learning's models as both *classifying* and *generative* correlate with the sociologist Pierre Bourdieu's conception of the *habitus*, a simultaneously cognitive and social "structuring structure" of "transposable dispositions" which unconsciously (and without symbolic rules) reproduces the stratifications of past experience (Bourdieu 1977). In addition, the increasingly baroque dyadic and/or dialogical (Bakhtin, 1981) architectures of the 'adversarial' generative networks of Goodfellow (2016), I argue, provide novel material and intellectual opportunities to rethink and reinscribe humanistic and social-scientific conceptions of agency and action *in vivo*.

Short bio
Michael Castelle is an Assistant Professor at the University of Warwick's Centre for Interdisciplinary Methodologies. He received a Ph.D. in Sociology from the University of Chicago in 2017 and also holds a Sc.B. in Computer Science from Brown University. His research combines the history and philosophy of science and technology with the economic sociology of digital markets, with influences from social theory and anthropological linguistics. His dissertation, "Transaction and Message: From Database to Marketplace, 1970-2000" traces the conceptual formalization and commercialization of both transaction processing and messaging middleware for the purposes of an increasingly digitized and globally interconnected financial industry.

Ansgar Koene, Horizon Digital Economy Research Institute, University of Nottingham, UK

Policy and regulatory approaches to AI systems – what role can ethical standards play
Despite warnings about Bias in Computer Systems going back to at least 1992, it has taken concerted efforts by groups of researchers like FAT/ML (Fairness, Accountability and Transparency in Machine Learning) and news and media reports highlighting discriminatory effects in data driven algorithms (e.g. recidivism prediction) to dispel the naïve-optimistic myth that the logic/mathematics of computation would automatically results in objectively unbiased outcomes. Starting from a brief overview of key challenges in the design and use of data driven algorithmic decision, as revealed by the youth- and stakeholder-engagement panels in our UnBias project (http://unbias.wp.horizon.ac.uk/), I will focus on the question of how these systems can be regulated to ensure that they contribute to improving human (and global) wellbeing. Against this background I will first discuss the IEEE Global Initiative on Ethics for Autonomous and Intelligent systems, highlighting the IEEE P7000 series ethics related standards that are being developed as part of the initiative. As illustrative example I will discuss the IEEE P7003 Standard for Algorithmic Bias Considerations. Finally, I will place these standards activities in the context of ongoing explorations regarding regulatory interventions that are being explored by the European Parliament and the European Commission.

Short bio
Ansgar Koene is a Senior Research Fellow at Horizon Digital Economy Research institute, University of Nottingham, and chairs the IEEE P7003 Standard for Algorithm Bias Considerations working group. Ansgar is lead researcher in charge of Horizon's Policy Impact; leads the stakeholder engagement activities of the EPSRC funded UnBias project to develop recommendations for minimizing unjustified bias in algorithmic systems; and is lead author for a European Parliament Science and Technology Options Assessment study on 'A governance framework for algorithmic accountability and

transparency'. Ansgar has published on topics ranging from AI, bio-inspired Robotics and Computational Neuroscience to experimental Human Behaviour/Perception studies.

Emanuele Nicolo Andreoli, School of Communication and Culture, Aarhus University, Denmark

POV and machine vision

The paper proposes to frame machine vision in the context of the proliferation of POV interfaces which are defining the aesthetic of a number of new technologies of vision (such as mobile phones, Google technologies of vision, virtual reality devices, and systems for machine vision), and which are turning POV into one of the most controversial political-aesthetic battlefields of our time. The paper argues that every process of grammatisation (Stiegler,1998) involves a question about POV, and within the broader attempt to elaborate an ontology and a phenomenology of the POV, it proposes to frame POV as the genetic element of any possible regime of visibility – also when it comes to the regimes of visibility produced by machine vision. In this context, the paper suggests to frame the notion of POV in relation to the rupture between the couple eye/ gaze (Lacan, 1964) and seeing/ seen (Merleau-Ponty, 1968) which machine vision seems to produce, and tries to define the meaning of POV in relation to a form of vision which bypass the isomorphism between technical image and perceptual image (Bellour, 1996) and produce images in absence of light (and of bodies) - indeed as assemblages of information (Shannon, 1948). To do so, the paper argues for the rehabilitation of the notion of the body as proposed by Hansen (2006) and for an understanding of the notion of information which moves away from Shannon's definition, and involves, instead, the presence of an embodied agent (Ruyer, 1954). The researcher plans to test this theoretical framework specifically in relation to so-called "dueling neural network" for deep reinforcement learning.

Sune Lehmann Jørgensen, Department of Applied Mathematics and Computer Science, Technical University of Denmark, Denmark

The role of data in making predictions

In the popular press, we often hear bold claims of how our personality, political leanings, or sexual preferences can be inferred based on "liking" seemingly unrelated items from our Facebook feeds. Or how our future mobility can be predicted based on a record of our past behavior. Likely, such examples are part of what causes public concern about the power of AI. In this talk, I discuss about the algorithms and data behind these to make the point that often, it's not the sophisticated AI algorithms that lead to spectacular findings as those mentioned above, but rather that such results are driven by a combination of boring/repetitive/stereotypical human behavior combined with the availability of rich behavioral data.

Short bio

Sune Lehmann is an associate professor at DTU Compute (Technical University of Denmark) and an adjunct (full) professor at the Department of Sociology (University of Copenhagen). Originally a physicist, Sune's work focuses on using analyzing and understanding large dataset arising from recordings of human behavior (e.g. GPS traces, cell phone calls, etc).

Christian Grund Sørensen, Department of Communication and Psychology, Aalborg University, Denmark

The Cartesian fallacy and the digital will

Since 1637 the "Cogito ergo sum" of Rene Descartes has been the primary touchstone of ontology. The cognitive rationality of man has been the philosophical sine qua non of human identity which is reflected not only in humanistic thinking but also in our view en technology. The "Turing-test" of Alan

Turing from 1950 thus evaluated a machine's level of intelligence by its ability to exhibit intelligent behavior equivalent to, or indistinguishable from, that of a human. When a machine displays a reasonable simulation of human utterance it is perceived as corresponding with human cognition and intelligence and thus, in the philosophical light of Descartes, becomes the bearer of some kind of anthropomorph existence. In the 21st Century, artificial intelligence has come very far in achieving Turing's dream. Neural networks, machine learning, exponential growth in computing power, and advanced algorithms have brought the Turing machine closer and closer to being the thinking machine. As a thinking machine, it becomes very close to human existence in Descartes definition. A less uncanny and often esthetically pleasant version of Mary Shelleys monster of Frankenstein (1818) with the notable difference that Shelley´'s monster displayed both intelligence and will whereas Turing's intelligent machine displays seemingly unlimited intelligence but cannot muster the will to tie a shoe lace. Perhaps Descartes' definition of human existence is highly problematic in the light of present technological advances. Perhaps the concept of will, and a discussion of deliberation in itself, is most relevant. The aim of this presentation is to discuss these issues.

Short bio

Born in Copenhagen in1968. M.Th. University of Copenhagen 1991. M.Sc. (Multimedia and worldviews) University of Aalborg 2007. Ph.d. (Worldview and Persuasion. Media, Munk, and Modernity) University of Aalborg 2016. Vicar in the Evangelical-Lutheran Church and researcher at The Kaj Munk Research Center, Aalborg University. Researcher in digital communication, cultural mediation through digital media, technology enhanced learning, and the Danish playwright and Pastor Kaj Munk. Lecturer in rhetoric, worldview, persuasive technology, information architecture, theology, philosophy, and ethics. Interest in working deeper with AI-issues and combining them with my competences in the areas of digital communication, philosophy and ethics.

Janet Rafner, Science at Home, Dept. of Physics and Astronomy, Aarhus University, Denmark

The Intuition Game: Hybrid Intelligence in fluid dynamics, quantum physics and other complex research problems

Almost daily news reports describe how AI has revolutionized yet another industrial domain. Some AI researchers believe that within a few decades it will surpass human capabilities in all domains, while others maintain that we are still far from understanding how humans reach fast, intuitive and correct decisions based on seemingly insufficient data. Thus, a huge challenge is developing ways to explore and model how humans use intuition and creativity to find or design solutions. In the www.scienceathome.org project, we have developed games allowing so far 250,000 players to contribute to research in quantum and classical physics, mathematics, chemistry, behavioral economics, corporate management, psychology and cognitive science. We believe that this wealth of data from individual and collective problem solving can be used to generate novel insights about human intuition and innovation that could potentially form the basis of novel forms of AI.Finally, I will describe our work within the new global educational movement, Think Like a Scientist, in which we introduce citizen science games to the formal school setting at all levels as a means to make the world of research and knowledge generation available to students from an early age.

Short bio

Ms. Rafner is an incoming doctoral researcher in hybrid intelligence at ScienceAtHome, Aarhus University. She was formerly a U.S. Fulbright Fellow, with degrees in physics and studio art. Her current research includes theoretical and phenomenological turbulence, Human Computer Interfaces, Research through Design, and Research-Enabling Game-Based Education. She is also the

international coordinator for the ScienceAtHome activities associated with the global educational initiative, Think Like a Scientist.

Kristoffer L. Nielbo, Datakube, Department of History, University of Southern Denmark, Denmark

It is just a machine that learns - on the role of computing and task automation in historical research

With the big surge in machine learning (ML), artificial intelligence (AI) is on everyone's lips. This is not different for research in the humanities, but here studies tend to focus on qualitative analysis and conceptual assessment of how, for instance, algorithms that learn from our data can reinforce prejudices, increase inequality, and fundamentally change humanity. Given such inhumane scenarios, it can be important to remember that underlying most AI is just rather "dumb" software that improves its performance on a pre-specified task with \experience." An ML-based program is however very useful for automating tedious and time-consuming tasks. For areas in social sciences and humanities that spend considerable time on indexing, curating and analyzing unstructured and soft data, such as history, this is good news. Important labor-intensive tasks involved in, for instance, source comparison, language annotation, and author profiling, can now be automated by machines, leaving only the executive decisions and interpretations to human computations. This presentation will discuss the impact of ML on the humanities with a particular focus on applications in historical research.

Anne Nielsen, Department of Clinical Medicine, Aarhus University, Denmark

Are convolutional neural networks able to predict tissue outcome in acute ischemic stroke?

Background and purpose: Each year, 13 million people worldwide suffers an acute ischemic stroke, making it a major disease and one of the leading causes of adult death and disability. Brain tissue infarcts permanently within hours after stroke onset and rapid treatment is therefore of utmost importance. Treatment options for acute ischemic stroke patients depend on the volume impacted by the stroke. Yet, this volume assessment is currently based on fixed thresholds, limiting accuracy. We wish to develop and validate a convolutional neural network (CNN) capable of automatically identifying the impacted tissue (final outcome) by utilizing the complex information derived from the acute MRI scan.

Methods: Using acute MRI, we developed and trained a deep CNN (CNNdeep) to predict the amount of impacted tissue. A total of N=187 patients were included. The performance of CNNdeep was compared to a shallow CNN (CNNshallow), a generalized linear model (GLM) and thresholding of the diffusion-weighted imaging biomarker apparent diffusion coefficient (ADC) at 600·10−6 mm2/s (ADCthres). The networks' performances were evaluated using visual inspection, area under the receiver operating characteristic curve (AUC) and contrast.

Results: CNNdeep yields significantly better performance in predicting final outcome (AUC=0.88±0.12) than GLM (AUC=0.78±0.12, p=0.005), and ADCthres (AUC=0.66±0.13,<0.0001) and a substantially better performance than CNNshallow (AUC=0.85±0.11, p=0.063). Measured by contrast, CNNdeep improves the predictions significantly, showing superiority to all other methods (p≤0.003).

Conclusions: The considerable prediction improvement accuracy shown by CNNdeep increases the potential for automated decision support in providing recommendations for personalized treatment plans.

Short bio

Anne completed her master in collaboration with Center of Functionally Integrative Neuroscience (CFIN), Department of Clinical Medicine, Aarhus University. Afterwards, Anne was employed as a Research Assistant at CFIN and Cercare Medical in October 2015. In the beginning of 2017, Anne started as an Industrial PhD student at CFIN and Cercare Medical. The project is entitled

"Convolutional Neural Network Based Prediction of Acute Stroke Response to Treatment" and focuses on the use of convolutional neural networks to predict the final outcome of acute ischemic stroke. The Industrial PhD project is funded jointly by Cercare Medical and Innovation Fund Denmark.

Jonathan Roberge and **Thomas Crosbie**, Institu Nationale de Recherche Scientifique, Canada

From Algorithmic Cultures to Machine Learning and Back

Recent literature on 'Algorithmic cultures' has been somewhat successful in combining elements of the cultural turn with STS, while attracting considerable attention —for instance, two Routledge editions (Seyfert & Roberge, 2016; Amoore & Piotukh 2016), a special Issue of *Science, Technology, & Human Values*, and a Conference at U.C. Berkeley (2016). It remains to be seen, however, how this theoretical framework could adapt to the latest trends in algorithmic knowledge production, especially in the field of neural network machine learning. As our focus thus changes, we must address specific new challenges, namely what constitutes *agency* and *governmentality* under these circumstances? *Agency* refers to the capacity to act purposively; for machine learners, it relates to the power of inferring and predicting outcomes (Mackenzie 2018; Burell 2016). New kinds of algorithmic control thus emerge at the junction of meaning-making and decision-making. The implications are tangible, particularly as machine learning becomes more unsupervised and begins to impact on work —not just repetitive tasks, but the broader responsibility belonging to what Reich once dubbed 'symbolic analysts' (1991). Connected to the complex entanglements of agency is the issue of *governmentality*, or how the field is organised, by whom, and for what purposes. Machine learning is based on an open-science model where public actors —governments and universities— are deeply entrenched, and where 'mathematical wonders' rapidly become consumer goods and services. However, while the algorithms themselves are openly available, the databases they are trained on are not. Hence the advantages for private actors such as Google or Facebook; hence the capability to build military AI with little regulatory oversight. In the end, and as our ethnographic work in Canada demonstrates, issues of *agency* and *governmentality* circle back to comprehensive ethical and cultural considerations.

Short bio

Jonathan Roberge, Thomas Crosbie and Kevin Morin respectively associate professor at INRS-UCS, assistant professor at Royal Danish Defence College and PhD student at INRS-UCS are a research team studying artificial intelligence and innovation ecosystem. Our research is funded by a grant from the Social Science and Humanities Council of Canada (2016-2018) and try to understand the effects of the rapid adoption of algorithmic technologies on everyday life. More specifically, we want to better understand the dissemination of increasingly predictive, automated and mobile, algorithmic data processing into mainstream consumer culture. We focus on fundamentals issues dealing with agency and governmentality which are changing in nature, especially when it comes to meaning-making and the production of an open-science field.

Anders Munk and **Anders Koed Madsen,** TANT lab and Aalborg University Copenhagen, Denmark)

Reasoning with algorithms: data sprints as experimental situations

Data intensive methods and algorithmic reasoning are currently at the center of attention in digital sociology. On the one hand, the potential of machine learning to solve hard problems that cannot, at present, be adequately theorized, is widely touted. On the other hand, concerns are raised about algorithmic bias and the risk of the machine uncritically and unreflexively reproducing the imbalances and asymmetries of social existence in undesirable ways. Both positions, although

dichotomous, fundamentally address the same question, namely the degree to which techniques for unsupervised learning are capable of accurately representing the world? Following the American pragmatist John Dewey and his concept of the 'experimental situation' we discuss how unsupervised learning can instead be turned into an object of inquiry in so-called 'data sprints' where it is the potential to raise new questions and slow down reasoning, rather than the ability to accurately represent, that becomes the critical parameter.

Duda Kvitsiani, DANDRITE - Kvitsiani Group, Aarhus University, Denmark

Distributed representations in cortical neural networks
Abstract TBA

Bendert Zevenbergen, Center for Information Technology Policy, Princeton University, USA

AI and Ethics Case Studies for Education
This presentation will showcase several case studies where the design and deployment of artificial intelligence systems was met by social or ethical dilemmas. The proposed analysis of ethical issues in artificial intelligence systems engineering will unpack some ethical considerations that guide decision-making in information systems or emerging technologies and apply them to the case studies. These case studies will have been published separately and for educational use across university departments and in professional training, in order to equip the next generation of engineers, managers, lawyers, and policy makers with a common set of intellectual tools when considering AI governance. The case studies are the result of a series of inter-disciplinary workshops held at Princeton University. Engineers from companies or government agencies that develop AI systems presented specific cases, the (social) problem to be solved, the technologies used, and the resulting moral dilemmas that arose during the development or deployment of the systems. The workshop participants consisted of about 30 scholars from computer science, data science, philosophy, social theory, and law. These participants scrutinized the case from moral, technical, legal, and other points of view and engaged in a structured discussion. The cases have been abstracted away from identifying features, but stay true to the presentation and discussion.

Short Bio
Ben Zevenbergen is a visiting professional specialist at CITP. His work mostly consists of multidisciplinary investigations in the ethical, social, and legal impacts of Internet technologies, and vice versa. His position is supported by Princeton's University Center for Human Values. He is currently finishing a Ph.D. at the Oxford Internet Institute about the research ethics for technical projects that involve unsuspecting Internet users as data subjects. Next to his doctoral work, Ben been working actively with computer scientists and network engineers to develop guidelines of ethics in networked systems research. Before returning to academia, Ben was a policy advisor to a politician in the European Parliament, working on Europe's Digital Agenda, copyright in international trade agreements and net neutrality. Previously, Ben worked as an ICT/IP lawyer and policy consultant in the Netherlands. He holds a degree in law, specializing in Information Law.