Autonomy in the Digital Age: Rethinking Relationships between Humans, Technology and Society

An international hybrid conference of the research group ‚Autonomy and Autonomous Systems'

20 November – 22 November 2022, University of Bonn, Germany & Online

University of Bonn, Lennéstraße 6 - Room 4.001, 53113 Bonn &

Universitätsclub Bonn e.V., Konviktstraße 9, 53113 Bonn

www.autonomy-research-group.org

- Abstracts -

Epistemic Shifts in the Digital Age: From Loving the Data to Loving Automation

Karin Knorr Cetina, O. Borchert Distinguished Service Professor, University of Chicago/USA

The way we work and live has long pointed to the role of professional knowledge in shaping the productive sector, our health care, the management of the economy, how we relate to the environment, the way we raise a child. The turn to science as a source of knowledge has been documented since the Enlightenment. But while we have all experienced the transformative power of science and technology in contemporary society, scientific knowledge-making itself also changes, it undergoes epistemic shifts of major and minor relevance. These shifts can be radical but they used to be internal to particular fields and disciplines, as suggested by the notion paradigm change or that of epistemic cultures. In this presentation, I begin to look at a shift we witness today in many fields, a shift toward an artificially intelligent science. I begin with the notion of "data" as a concept that captures empirical reality but I will also discuss the notion automation and its replacement by autonomy as well tensions that arise around established scientific codes and the challenges that occur when representational vocabularies and our methods of sustaining them become challenged.

Europe in a Geo-economic Age: Digital Sovereignty or Digital Dependence

Abraham Newman, Professor Edmund A. Walsh School of Foreign Service and Government Department, Georgetown University/USA

For Europe, economic openness long served as its primary path to wealth and power. Increasingly, however, interdependence has become a source of vulnerability as well as autonomy. Shocks like the COVID-19 pandemic, exposed world-wide economic networks for information, money, and production as more brittle and centralized than many observers had expected. But this is far from simply an economic story as the US and China have begun to manipulate these vulnerabilities, inflicting costs on their adversaries as a tool of foreign policy. The Russia invasion of Ukraine and its gas war has only made this more obvious. The chokepoints of the global economy have become the battlefield upon which interstate tensions playout. Unfortunately, Europe increasingly finds itself stuck in the middle. The objective of the talk is to map how these vulnerabilities have emerged and how they impinge on European autonomy and what it might do to navigate this new era of economic coercion.

Autonomy of the digital subject

Gabriele Gramelsberger, Professor Theory of Science and Technology, Human Technology Center, RWTH Aachen/Germany
Maximilian Mayer, Junior-Professor International relations and global technology policy, University Bonn/Germany

The rise of digital authoritarianism—its techniques, laws and politics—coproduce a new type of individual, one less autonomous than democratic theory and practice previously demanded. The key question to ask is what kind of citizen/consumer/user is becoming normalized in the digitalized world (proposing the metaverse). In particular two aspects are of interest: Firstly, the phenomenological condition of the digital subject, and secondly, the impact of these conditions on the democratic subject. In this paper we explore from a politological as well as philosophical perspective how these sociotechnical changes reshape the individual and the scales at which they have consequences for democratic life. We will ask what kinds of subject are being coproduced alongside with authoritarian digitalization?

On different uses of the concept "autonomy" & Autonomy, Technology, and Artificial Intelligence

Dieter Sturma, Professor Philosophy with special reference to ethics in the life sciences, University Bonn/Germany
Bert Heinrichs, Professor Ethics and Applied Ethics, University Bonn/Germany

Autonomy, Technology, and Artificial Intelligence, *Dieter Sturma*

The concepts of intelligence and autonomy are not defined clearly. However, it can be stated that only persons are sources of initiatives, accountability, and consideration. This distinguishes them from all technical systems. The narrow sense of the term "autonomous" is contrasted by a linguistic practice which applies to technical systems as well. This affects both the way people conceptualize themselves and their scope for action.

In the history of the human life form, technology has successively expanded the possible scope of action—first with and word, much later with the help of technical devices and so-called autonomous systems. From the start, technology has been an integral part of the human space of reasons with impacts in the space of causes. Therefore, the path that eventually leads to AI goes back much further than is commonly assumed. It is very probable that technology in its first forms did not manifest itself as pure means or tools, but as regulated movement or rule-governed behaviour—inspired by observation of natural processes. The primal scene (*Urszene*) of technology is in this respect not a rupture, but a change in the course of events with new means. These beginnings occur as complex behavioral and mental transitions with an obviously high potential of self-dynamics. These self-dynamics is one cause of technically induced destructions in the social and natural environment.

The artificial worlds of technology have an effect on humans and blur the outlines of their presumed nature. Due to the epistemic and practical intertwining of nature and culture, the human form of life is always located beyond naturalness. Technology takes the form of an extended hand—axe, a hatchet, sword, spade, or rope—as a learnable ability or skill—language, arithmetic, music, building, or metalworking—as an independent tool—ship, plow, wheel, wagon, or winch—and as automation—mill, loom, steam engine, motor, artificial intelligence, and robotics.

Forms of technology are always connected with activities that are equally determined by human persons and technology. Technology functions only under the conditions of a "cooperative" nature. Persons simply need connecting points for their technical interventions, which, to be effective, must extend into nature. This results in mutual constraints. Humans must manipulate objects or resources in such a way that the respective technical applications can operate and produce the expected outputs. In doing so, however, they must adapt to mechanisms of application—to cutting tools, levers, and weights, for example, but also to reduced sensor technology or specified algorithms. While technology is overtly recognizable as a manifestation of the mind, it is easily overlooked that people are constrained by technical specifications in their epistemic and practical attitudes.

Regarding artificial intelligence, the dream of strong artificial intelligence should give way to specified projects of expanding the scope of human activities. These expansions will result in shifts of the human-machine interface. Moreover, cultural goals need to be identified as means of which AI can operate—explorations of outer space being a good example for this. A close cooperation of AI research and normative evaluation is suitable to establish technical systems as tools of a cultural process in the sense of personal and social autonomy. In this way, so-called autonomous systems are always possible functions of the autonomy of persons. It is therefore not so much about the mechanization of the social world, but rather about the contribution that artificial intelligence makes or can make to a *justified* humanizing of the life world (*Humanisierung der Lebenswelt*).

Cantwell Smith, Brian: *The Promise of Artificial Intelligence: Reckoning and Judgment*, Cambridge, Mass.: MIT Press, 2019.
Cassirer, Ernst: „Form und Technik", in: *Symbol, Technik, Sprache*, Hamburg, 1985.
Leroi-Gourhan, André: *La geste et la parole II, La mémoire et les rytmes*, Paris, 1964.
Sellars, Wilfrid: "Languages, Rules and Behavior," in Jeffrey F. Sicha (ed.), *Pure Pragmatics and Possible Worlds. The Early Essays of Wilfrid Sellars*, Atascadero: Ridgeview, 1980.
Sellars, Wilfrid: "Philosophy and the Scientific Image of Man," in: *Science, Perception and Reality*, Atascadero 1963, 38–40.
Sellars, Wilfrid: *Empiricism and the Philosophy of Mind*, Cambridge, Mass., 1997,
Sturma, Dieter: Persons. A Thick Description of the Human Life Form, in: Jörg Noller (Hg.): *Was sind und wie existieren Personen?*, Leiden 2019, S. 147–165.

On different uses of the concept "autonomy", *Bert Heinrichs*

Concepts are never unambiguous, neither in every-day language nor in professional discourse. While it is often sufficient to indicate different usages, occasionally they lead to serious irritation or even dispute. A case in point is the concept of autonomy. In the humanities, the concept has a long tradition and is intimately connected with human self-understanding; in computer science, on the other hand, it is being used to denote artificial systems with special properties. The gap between these two usages is significant. For the very fact that artificial systems are called "autonomous" threatens to undermine the idea of a "special status" [*Sonderstellung*] that the concept was traditionally intended to indicate in humans. Following this diagnosis, different uses of the concept point to a deeper question, namely whether the talk of a "special status" of humans is still tenable. In this paper I shall argue, on the one hand, that assuming a special status of humans is no longer convincing. In particular, autonomy should not be understood as some sort of mysterious capacity, but rather as one that artificial systems could in principle possess. On the other hand, however, the traditional understanding of autonomy can be translated to mean the practice of giving and asking for reasons. Even the most advanced computer systems are not yet capable of this discursive practice. In this respect, autonomy does mark a certain

(albeit contingent) special position of humans. Different uses of the concept of autonomy should not obscure this.

Brandom, R. B. (1994). *Making it explicit. Reasoning, representing, and discursive commitment.* Harvard University Press.

Heinrichs, B., Knell, S. (2021). Aliens in the Space of Reasons? On the Interaction Between Humans and Artificial Intelligent Agents. *Philosophy and Technology* 34, 1569–1580. https://doi.org/10.1007/s13347-021-00475-2.

Sellars, W. (2007). Philosophy and the scientific image of man. In K. Scharp & R. B. Brandom (Eds.), *In the space of reasons. Selected essays by Wilfrid Sellars* (pp. 369–408). Harvard University Press.

Farewell, Traffic Law

Guy Seidman, Professor Harry Radzyner Law School, Reichman University, Herzliya, Israel

Dr. Aviv Gaon, Senior Lecturer Harry Radzyner Law School, Reichman University, Herzliya, Israel

We know that no law student has ever entered law-school with the hope of specializing in traffic law. Few if any law schools even offer such a course. Some might say that there is no need for traffic law – like there was no 'law of the horse' (and no 'internet law'). But this is only to say that humans centered legal disciplines around themselves – human acts, and omissions, human ownership, and discretion. Traffic has dramatic legal effects. Human driven vehicles cause enormous harm. The hazards of human driving, and the social concerns they raise, i.e., traffic law, are handled dealt by the better-defined fields of insurance, tort, and criminal law. One of the possible effects of the age of autonomous vehicles (AVs), especially if, in congested urban areas, human driven cars will be banned, is the decline of traffic law. In weighing the pros and cons for a Future Without Human Driving (and drivers), we consider the shaded ground of the social implications of driving. Taking aside, for a moment, the other essential elements of driving we considered in other parts, and the likely results of autonomous vehicles (AVs) arrival and regulation, one should also take the social impacts in mind. Perhaps the greatest promise of AVs is their inter-connectivity: traffic will no longer start at the pace of human reflexes, cruise at a speed decreed by each individual drivers' discretion or change course at the good-will of fellow drivers, noting car signaling. Better coordination – and much less human interaction and discretion – is likely to eliminate human fault and error from the equation. Leading, so we hope, to efficient traffic, fitting road and weather conditions, equitably applying to all drivers. If all goes well – and does it ever? – than traffic law will much decline, if not outright disappear. Relieving human drivers of their roles as drivers will likely reduce much of the stress that many humans report on their daily routine; however – what will be the expected outcome of this change in human behaviour, and how will this change affect our legal norms and standards? Perhaps carpooling would now become a more obvious social event? Perhaps we should encourage larger Avs, rather than plan for smaller, one-person pods? Perhaps we should consider using the resources freed by Avs – fewer roads, much less parking – specifically for encouraging more socially-minded goals – such as communal recreational parts, sports and culture facilities? Or maybe we should actively encourage human driving on traffic-free days, such as the weekend? Those questions are all policy-related questions that will inadvertently affect AVs regulation. In urban settings, this is likely to lead to dramatic changes. Consider, for example, the possibility of eliminating all traffic signals and traffic lights – human eyes to be replaced by electronic communications and sensors; consider the moneys saved and light pollution reduced as cars need no lighting to see where they are going; consider the lives and limbs saved with fewer accidents and the

legal expenses saved with fewer traffic tickets and subsequent litigation. Having said this, there is another side to it. Surely there is little to be said in favor of DUI, but the elimination from human discretion in driving will diminish driving as a human experience. Counterintuitive as this may sound, there are societal benefits for maintaining the road as a communal meeting place, perhaps the last remaining.

Listening Devices. Music Media of the Pre-digital Era

Jens Papenburg, Professor Sound Studies, University Bonn/Germany

Since the invention of the phonograph and the gramophone in the late 19th century listening (to music) has been organized more and more in relation to technology and devices. Although such devices were and often are easily accessible, up to now we have no elaborated concept of them. To address this gap, I propose in my presentation the term "listening device". In conjunction with this concept, I explore a fruitful method for analysing listening as a historical subject that has been increasingly organized in relation to technology.

The main objective of my presentation is to develop "listening device" as a concept relying on sound studies, media studies and cultural theory. Despite all the characteristics specific to the different listening devices, they can nevertheless be compared because of the fundamental similarities they share: they model and manage listening, they actively mediate between the listener and the music heard, and it is this mediation that brings both listener and the music listened to into being. Ultimately, however, the intention is that the listening devices themselves should not be heard so that the music they playback can be heard. Thus, listening devices take the history of listening to its very limits and confront it with its "other"-a history of non-listening.

My presentation deals primarily with listening devices of the pre-digital era. However, in the conclusion I try to give an outlook how these considerations can be developed for an analysis of listening in the digital era and how listening devices turned into more and more autonomous listening machines.

The Digital Revolution and Democratic Politics: A Problem-Solving Divergence

Peter Hägel, Associate Professor for Philosophy, Politics & Economics, The American University of Paris/France

I propose a novel hypothesis within the discussion of declining trust, anti-politics and democracy, probing the OECD's surface observation that "[g]overnments are operating in an increasingly complex and fast-paced environment with growing expectations from citizens for an efficient and seamless interaction with their governments" (2022: 67). Over the past 30 years, innovations in digital technologies and related business practices (e-commerce, customer service, personalization, logistics, etc.) have tremendously augmented people's problem-solving capabilities in daily life. When we face a problem, online retailers and service providers deliver solutions, or we visit apps and websites where experts and peers provide pertinent DIY-instructions, often in real-time. Much more than in the past,

we find and get what we want, almost instantly. The resulting experiences and accompanying narratives, within a broader socio-cultural context of "life hacking", "self-optimization" and "instant gratification" (Reagle 2019), create expectations about problem-solving that lead to a growing divergence between daily life and democratic politics. At the input (will formation), throughput (decision-making process) and output (public service provision) levels, democratic politics remains Weber's "slow, powerful drilling through hard boards". But the complicated art of coalition-building and compromise-making (Stoker 2017) is more and more at odds with people's everyday experience, and therefore appears as unapt and outmoded.

OECD (2022). Building Trust to Reinforce Democracy: Main Findings from the 2021 OECD Survey on Drivers of Trust in Public Institutions. Paris: OECD.

Reagle, Joseph M. (2019). Hacking Life: Systematized Living and its Discontents. Cambridge: MIT Press.

Stoker, Gerry (2017). Why Politics Matters: Making Democracy Work. 2 nd ed. London: Palgrave

Limiting technologisation for human autonomy? Reconsidering Illich's "conviviality" in the digital age

Anne Hoss, Doctoral research assistant at the Chair of Political Science, Helmut Schmidt University of Hamburg, Hamburg, Germany

The proposed presentation aims to contribute Ivan Illich's (1973)1 concept of socio-technological "conviviality", which suggests that high-tech development and human self-determination are only compatible to a limited extent. His work from the heyday of technology critique in the 1970s considers modern technology's energy needs, its economic imprint and related power structures as fundamentally detrimental to realising individual and collective autonomy. These heteronomising features, Illich argues, can only be met by a reorientation of innovation policy, but also a (deliberatively defined) limitation of high-tech diffusion. With this radical approach, Illich goes beyond most current perspectives that emphasise potentials of shaping continual digitalisation processes according to social needs. The presentation seeks to discuss Illich's ideas in the light of present debates on autonomy in the digital age. It will argue that despite limits of applicability, Illich's "thick" concept of autonomy offers a demanding, debatable normative framework for reflections on the meaning, possibilities and conditions of self-determination in our digital societies.

"Autonomous weapons" as a geopolitical signifier in a national power play: analysing AI imaginaries in Chinese and US military policies

Thomas Christian Bächle, Alexander von Humboldt Institute for Internet and Society, Germany

Jascha Bareis, Institute for Technology Assessment and Systems Analysis (ITAS) Karlsruhe Institute of Technology, Karlsruhe, Germany

“Autonomous weapon systems” (AWS) have been subject to intense discussions for years. Numerous political, academic and legal actors are debating their consequences, with many calling for strict regulation or even a global ban. Surprisingly, it often remains unclear which technologies the term AWS refers to and also in what sense these systems can be characterised as autonomous at all. Despite being feared by many, weapons that are completely self-governing and beyond human control are more of a conceptual possibility than an actual military reality. As will be argued, the conflicting interpretations of AWS are largely the result of the diverse meanings that are constructed in political discourses. These interpretations convert specific understandings of AI into strategic assets and consequently hinder the establishment of common ethical standards and legal regulations. In particular, this article looks at the publicly available military AI strategies and position papers by China and the USA. It analyses how AWS technologies, understood as evoking sociotechnical imaginaries, are politicised to serve particular national interests. The article presents the current theoretical debate, which has sought to find a functional definition of AWS that is sufficiently unambiguous for regulatory or military contexts. Approaching AWS as a phenomenon that is embedded in a particular sociotechnical imaginary, however, fags up the ways in which nation states portray themselves as part of a global AI race, competing over economic, military and geopolitical advantages. Nation states do not just enforce their geopolitical ambitions through a fierce realpolitik rhetoric but also play around with ambiguities in definitions. This especially holds true for China and the USA, since they are regarded and regard themselves as hegemonic antagonists, presenting competing self-conceptions that are apparent in their histories, political doctrines and identities. The way they showcase their AI-driven military prowess indicates an ambivalent rhetoric of legal sobriety, tech-regulation and aggressive national dominance. AWS take on the role of signifiers that are employed to foster political legitimacy or to spark deliberate confusion and deterrence.

A feminist lens on automation in education

Irina Zakharova, Postdoctoral researcher on datafication of education and feminist data studies at the university of Bremen, Bremen, Germany

Juliane Jarke, Professor for Digital Society at the Graz University, Graz, Germany

Dataficaiton of education - a translation of teaching, learning, organisational, and governance processes into digital data poses new questions about autonomy and automation in education. As a backbone of democratic societies raising citizens to participate in social, political, and economic life, educational domain requires particular scholarly attention. Digital data make automation of education possible, and educational technology providers promise their customers more space, time, and ultimately autonomy for individual support and creative engagement with students. In educational practice, this promise of automation and autonomy, however, often cannot be accomplished, as presumably 'autonomous' systems require a lot of attention and labour from their human users to function.

In this polarized context of what automation is and what benefits and harms can it bring forth, we are interested in the intricate ways in which both liberating and debilitating aspects of automation co-exist and how educational actors find ways to cope with these conflicting positions. Specifically, drawing on

feminist ethics of care (Tronto, 1993) we discuss various shifts of power in contexts of different datafied and automated processes in schools. We report from several broader empirical projects on datafication of K-12 school education in which we interviewed educational technology providers, teachers, and school management. We contribute to research on automated systems conceptually and empirically by discussing modes of autonomy these systems allow in schools.

Tronto, J. C. (1993). *Moral boundaries: A political argument for an ethic of care*. Routledge

Working With Robots: Design and Evaluation of an Introductory Computer Science Teaching Unit With Educational Robots

Lilli Bruckschen, Postdoctoral researcher at the Humanoid Robots Lab at the University of Bonn/Germany

As our world continues to digitalize more and more, Computer Science concepts have started to interweave with our daily life. Accordingly, teaching these concepts in schools is becoming increasingly relevant. An illustrative and practical way to do this is by using haptic examples of these very same concepts in form of educational robots. This offers the benefit of motivating and playful access to the field for young students. However, to integrate the robots productively into Computer Science lessons, engaging teaching units are essential. To support the design of those teaching units, we surveyed students and teachers to evaluate their preferences regarding the use of robots in Computer Science lessons. The survey had 95 participants, 6 teachers, and 86 students, from 6 different classes of 4 different schools. Using the results of this survey, we further designed, conducted, and evaluated a teaching sequence for a German 6th-grade Computer Science course. As a robot, we used the educational kit LEGO MINDSTORMS (Model 51515) alongside Apple iPads. The course consisted of 30 students, 15 of them male and 15 female, aged between 11 and 13. During the sequence, we observed the students' motivation and their progress in learning computer programming concepts. We also examined the results of their exercises. We found that the robots positively impacted the students' motivation and learning process.

Chernobyl 1986/2086: Ripping the Textile of Time Woven into Artificial Intelligence

Benjamin Peters, Fellow Käte Hamburger Kolleg RWTH Aachen University, Aachen, Germany

The present ripped on April 27, 1986. Like artificial intelligence (AI), Chernobyl, once the future of a nuclear, electrified Soviet socialism promised a more powerful, prosperous, and peaceful future that stretched into forever until, unlike AI today, it was no more. In this draft chapter of my project on Soviet AI, I draw on new Ukrainian-language archives and interviews with cybernetic roboticists on site in 1986 and the 1990s International Space Station cases to challenge future-oriented imaginings (both

apocalyptic and "eucatastrophic") of artificial intelligence. I frame my alternative genealogy of smart tech around the proverbial metaphor of the present as a felt fabric of the present, a past filtered forward. By stretching and restitching tears in the fabric of Chernobyl and contemporary AI discourse, I hope to model a new way to talk about the textile of time woven into AI.

Designing Agents as Artificial Companions - A Systematic Literature Review on the Companion Paradigm

Ayanda Rogge, Research associate at the Institute of Media and Communication, TU Dresden, Dresden, Germany

Recurring terms like friend or companion describe technological agents in social contexts. We investigated the concept behind these ascriptions: the companion paradigm, which represents a design principle to build agents that appear as artificial companions (ACs) to their users and form human-machine relationships [1, 2]. Although ACs seem to be a popular research subject, it is not clear what distinguishes an AC from other agents not claimed as such. Hence, we see the problem of a missing common denominator for what an AC is and what this implies for the agent's design. To approach this gap, we conducted a systematic literature review (SLR) of n=540 AC papers published since 2000 to gather defining attributes and synthesize AC characteristics. We assume that how an agent communicates and interacts with its user crucially shapes their perception as an AC. Accordingly, our findings suggest the companion paradigm as a design principle of interest for any assistive technology interacting with a human user. Furthermore, as ACs aim for highly personalized interaction and emotional engagement, we see the chance to create resilient human-machine teams. This can be relevant in many social areas, such as care, education, personal assistance or where users are socially isolated. Our SLR contributes to existing AC research by suggesting a definition (RQ1) that is systematically derived and integrates interdisciplinary research efforts on ACs. Also, it provides a list of AC characteristics (RQ2) structured in eight groups: verbal and nonverbal communication, adaptivity to user and context, engagement, personality, presence, and multimodality. Additionally, based on the performed SLR, we propose an AC typology (RQ3) to distinguish ACs regarding their application, user, task, and role.

Biundo, S., & Wendemuth, A. (Eds.). (2017). Companion Technology. Springer International Publishing. https://doi.org/10.1007/978-3-319-43665-4

Böhle, K., Bopp, K.: What a vision: The artificial companion. A piece of vision assessment including an expert survey. Science, Technology & Innovation Studies 10(1), 155–186 (2014

Human-Autonomy Teamwork (HAT) on Construction sites: Agency, Communication, and Collaboration

Sven Stumm, Chair for Individualized Production RWTH Aachen University, Aachen, Germany

Caja Thimm, Professorship for Media Studies and Intermediality, University Bonn, Bonn, Germany

Sigrid Brell-Cokcan, Director of the Chair of Individualized Production at RWTH Aachen University, Aachen, Germany

The demand for technical assistance at construction sites is growing, due to the increasing shortage of skilled and experienced construction workers[1] and the rising demands for specialized work in the construction industry. Construction companies are therefore increasingly relying on automation solutions and knowledge management systems to secure skilled workers and retain expertise (Polzin et. al., 2022).

As a consequence, many processes on construction sites rely on high- performance cooperation, which is more and more determined by human-machine teamwork. Research on 'human–autonomy teamwork' (HAT) promises a fruitful approach to these challenges.

Human–autonomy teamwork involves humans working interdependently toward a common goal along with autonomous agents. HAT has also been described as at least one human working cooperatively with at least one autonomous agent (McNeese et al., 2018), where an autonomous agent is a computer entity with a partial or high degree of self- governance with respect to decision- making, adaptation, and communication. Autonomous agents also involve a degree of self- directed behavior (agency) They take on a unique role or set of tasks and work interdependently with human team members to achieve a shared objective (O'Neill et al., 2022).

But the teaming process is not only depending on the degree of attributed autonomy and effectiveness to the machine, but also on human attitudes and experiences. Studies show that in order to understand better how humans accept autonomous agents as teammates rather than tools, factors like involved agency, benevolence, communication, interdependence, and coordination have be taken into account (Lyons et al., 2021).

For the context in question, the paper will explore the interdisciplinary options and challenges of HAT in the respective context of team work on the construction site. In the presentation, we will base our reflections on a critical assessment of the concept of 'autonomy levels' for the HAT approach and look at practical challenges of HAT in the context of human group perceptions and agency on the part of the human and the machines.

Lyons, J.B., Sycara, K., Lesi, M., Capiola, A. (2021). Human-Autonomy Teaming: Definitions, Debates, and Directions. *Frontiers in Psychology 12*.

McNeese, N., Demir, M., Cooke, N., Myers, C.W. (2018). Teaming with a synthetic teammate: insights into human-autonomy teaming. *Human Factors 60*, 262-272.

Polzin, B., Ringler, P., & Weigl, H. (2022). Führungsaufgabe Wissensmanagement. In Wissensmanagement im Bauwesen (pp. 25-31). Springer Vieweg, Wiesbaden.

O'Neill, T., McNeese, N.J, Barron, A, Schelble, B. (2022). Human–Autonomy Teaming: A Review and Analysis of the Empirical Literature. *Human Factors,* Vol. 64(5) 904–938. DOI:10.1177/0018720820960865

[1] https://de.statista.com/infografik/17357/fachkraeftemangel-nach-branchen/ last visited 10/18/2022

Autonomous machines as members of innovation teams: Prompt Engineering as a novel collaboration skill to bridge between human and machine intelligence in engineering and design

Frank Piller, Professor of technology and innovation management at RWTH Aachen University, Aachen, Germany and Academic Director of the Institute for Management Cybernetics

Kathleen Diener, Professor of Business Information Systems & Digital Innovation, Niederrhein University of Applied Sciences and Scientific Director of the Affiliated Institute for Business Cybernetics e.V.

When algorithms (a machine intelligence) with a high degree of autonomy take on entirely new roles in the management context and become the decision maker, our fundamental understanding of how to organize the division of labor changes completely. Until now, there has been a clear separation between tasks performed by machines and tasks performed by humans. As long as enough training data is available, machine intelligence can handle unstructured data (text, images, and video) well. However, it cannot fully represent complex business problems in organizational contexts or situations characterized by fuzzy decision rules (Dellermann et al., 2019; Gigerenzer, 2022). Machine intelligence is also poor at solving multiple tasks simultaneously (Raj & Seamans, 2019).

This suggests a combination of human and machine intelligence capabilities. Dellermann et al. (2019) propose the term "hybrid intelligence" for this. The goal is to use the complementary strengths of human and machine intelligence in such a way that a better overall performance can be achieved than if machines or humans were to decide alone. Thus, human intelligence and machine intelligence adapt and collaborate, creating a two-way exchange and mutual control. Thus, the concept goes beyond previous discourses on AI-augmented decision-making, which assume a sequential separation of human and machine contributions. Machine intelligence can also take on the role of a work partner or even a supervisor, as suggested in the debate on algorithmic management pre (Lee et al., 2015). In a working system, humans and AI need not fight each other but can complement each other. Autonomous decision-making agents (algorithms) become teammates (Seeber et al., 2020).

Hybrid intelligence requires rethinking a fundamental economic issue: the design of the division of labor and distribution of tasks in an organization. While computer science designs decision models and data structures and engineering the corresponding technical applications of AI, implementing hybrid intelligence is an economic (management) phenomenon (von Krogh, 2018). From the perspective of decision making theory, this raises the question of how we can balance the degree of autonomy in decision-making between human and algorithmic players, i.e., how and when would we assign decision-making authority to a human, a machine intelligence, or a participatory setting of a hybrid intelligence? From a conceptual dimension, this research asks how to rethink autonomy in a management context in the age of hybrid intelligence, when the traditional boundaries between human and machine decision-makers are blurred in interdisciplinary teams of a new order.

A concrete use case of hybrid intelligence is the integration of AI into R&D processes, i.e., hybrid decisions in the environment of creative processes and open problems. In contrast to AI integration in operational processes in production and logistics or administration (Robotic Process Automation), the integration of autonomous decision systems in R&D is an emerging application domain (Piller et al., 2022; Verganti et al., 2020). In this context, the concept of *Prompt Engineering* recently evolved as a way humans work with complex AI systems, particularly natural language processing (NLP) tools like GPT-3. The more autonomous machines become in providing the desired outcome, e.g., a new product

concept or a set of functional requirements for a new design, the more critical the task that is promoted to the machine producing the outcome. In some recent accounts, the role of the human engineer of the future is seen in designing the prompts while the machine does the rest. The presentation will reflect on the rise of hybrid intelligence in design & engineering, focusing on recent developments in prompt engineering.

Strategic Autonomies: Some reflections on the politics of autonomous systems

Lucy Suchman, Professor Sociology (em.), Lancaster University/England

In this talk I draw from work within Science and Technology Studies (STS) that has challenged, and reconfigured, dominant conceptions of autonomy as an attribute of either persons or technological systems. From Langdon Winner's writings in the 1970s to critical work in the present moment with respect to artificial intelligence and robotics, STS has been committed to empirically-based investigations that refigure autonomy from an inherent capacity to an effect of specifically situated discourses, political economies, and material practices. I trace the implications of this scholarship for my own current work on developments in the further automation of military systems, including the tensions and contradictions that haunt the call for meaningful human control. I conclude with some thoughts about how the figures of closed and open worlds can help us to delineate the limits of automation and to challenge investments that sustain the fantasy of autonomous machines.

Handing over Control to Autonomous Systems and Robotics in Construction the Strain of Working with Complex Systems

Sven Stumm, Chair for Individualized Production RWTH Aachen University, Aachen, Germany

Sigrid Brell-Cokcan, Director of the Chair of Individualized Production at RWTH Aachen University, Aachen, Germany

Compared to other industries, the level of automation in the construction industry is currently still low. A central reason for this is the incapability of current technologies to autonomously perform complex tasks in dynamic and unstructured environments such as construction sites. In these work environments the collaboration between humans and robots is key in order to perform construction tasks successfully that go beyond inspection and surface processing and require direct physical interaction with the construction environment. Previous research in the field of Human Robot Collaboration (HRC) has shown the complexity of the relationship between worker and machine in these interaction (Johnson et. al., 2009). The work of Shayesteh et. al. (2021) has shown that working with autonomous collaborative robots can induce a significantly higher cognitive load for construction workers than semi-autonomous collaborative robots. Handing over control, the concern associated

with this, as well as miscommunication between worker and robot have been stated as major contributors to this higher cognitive load. Within our work we take a look at the different levels and methods of communication in HRC, as well as potential benefits that robotics holds as a communicator in the mediation between digital planning and execution of construction tasks on-site (Stumm, 2019).

Johnson, R. C., Saboe, K. N., Prewett, M. S., Coovert, M. D., & Elliott, L. R. (2009). Autonomy and automation reliability in human-robot interaction: A qualitative review. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 53, No. 18, pp. 1398-1402). Sage CA: Los Angeles, CA: SAGE Publications.

Shayesteh, S., & Jebelli, H. (2021). Investigating the Impact of Construction Robots Autonomy Level on Workers' Cognitive Load. In Canadian Society of Civil Engineering Annual Conference (pp. 255-267). Springer, Singapore.

Stumm, S. (2019). Interconnecting design knowledge and construction by utilizing adaptability and configurability in robotics: mediating digital information from architectural design to construction through parametric design intent based robot programming (No. RWTH-2019-07757). Lehrstuhl für Individualisierte Bauproduktion.

The Mode of Autonomy – AWS and the doning of Iran Air Flight 655

Christoph Ernst, Research associate Media Studies, University Bonn/Germany

A key question in the discussion about autonomous weapon systems is the morally and ethically difficult question of 'meaningful human control' of these systems. Against this backdrop, the talk addresses a specific problem of media studies research on autonomous weapon systems. The focus is on the knowledge that exists about the 'operational mode' of a weapon system. Using the example of the downing of Iran Air Flight 655 in 1988, it will be shown that is only possible to talk in an analytical sense about 'meaningful human control' if the human-machine-interfaces of autonomous systems and their specific logic is taken into account.

Negotiating Normativity: Practices, technology, and lethal autonomous weapon systems

Ingvild Bode, Associate Professor at the Centre for War Studies, University of Southern Denmark, Odense, Denmark

Lethal autonomous weapon systems (LAWS) are the subject of considerable international debate turning around the extent to which humans remain in control over using force. But stakeholders disagree about how to understand the basic technologies that animate LAWS. Automation, autonomy, and AI are all common terms used, but they trigger diverging connotations. Such differences matter because they shape the substance of the debate and which regulatory options are put on the table. I therefore ask: How do discursively performed practices shape normativity in the international debate about LAWS? To understand this process, I draw on practice theories, science and technology studies (STS), and critical norm research. I argue that a constellation of communities of practice (Adler 2019) shapes the public debate about LAWS and focus on three of these communities of practice: diplomats, weapon manufacturers, and journalists. Actors in these communities of practice discursively perform practices of boundary work, in the STS sense, a communally approved form of drawing the lines between science and politics to shape understandings of automation, autonomy, and AI (Jasanoff 1987; Suchman 2012). These practices of boundary-work are normative in drawing implicitly or explicitly on ideas of oughtness and justice. Using data collected across four qualitative methods (e.g. participant observation, expert interviews), I analyse these dynamics empirically in two steps: first, I focus on how practices in relation to defining automation, autonomy, and AI performed by diplomatic stakeholders at the Group of Governmental Experts (GGE) in Geneva, the institutional focal point of this debate, shape normativity at a general level. Second, I study how this 'terminology à la carte' plays out in

discursive practices performed in relation to the UN-reported use of Kargu-2 loitering munitions, a particular type of weapon systems with autonomous features in targeting, in the Second Libyan Civil War.

Adler, Emanuel. 2019. World Ordering: A Social Theory of Cognitive Evolution. Cambridge: Cambridge University Press.

Jasanoff, Sheila S. 1987. "Contested Boundaries in Policy-Relevant Science." Social Studies of Science 17 (2): 195–230.

Suchman, Lucy. 2012. "Configuration." In Inventive Methods. The Happening of the Social, edited by Celia Lury and Nina Wakeford, 48–60. London & New York: Routledge.

From alignment to adaptation. Histories of automation in robotic radiosurgery

Kathrin Friedrich, Professor Digital Media Culture, University Bonn/Germany

Probably the most heard commands in medical practice still are- "Hold your breath", "Breath normally" or "Hold still"/"Don't move". To examine and treat the living body it needs some kind of regularity and regulation and some kind of stillness to be observable by a trained and embodied ear, eye, or hand. This not only holds true for diagnostics but also for current interventional practices where digital technologies play a fundamental – if not exclusive – role for planning and controlling, for example, surgical interventions.

Robotic radiation surgery also fundamentally relies on an array of media technologies as well as new forms of collaboration between physicians, patients and robotic actors. This collaboration entails to intraoperatively adapt the moving patient's body to surgical technologies. In robotic radiosurgery it is literally vital to remotely synchronize the body with the automated targeting of interventional radiation beams that are aimed to destroy pathological tissue inside the body. If the predetermined target gets out of the center of radiation, too much healthy tissue might be destroyed and cause unwanted after effects.

The paper discusses the mutual and vital adaptation of bodies and intervention in robotic radiosurgery focusing on histories of automation within the field. By tracing historical and current techniques of targeted radiation it will show how the quite material alignment of patients to surgical procedures were transformed into a matter of adaptation through the means of digital technologies.

Undue Manipulation or Autonomy Support: A Cross Section of the Ethics of Clinical Decision Support-Systems

Hendrik Kempt, Postdoctoral researcher, Human Technology Center, RWTH Aachen/Germany

Jared Parmer, Postdoctoral researcher, Human Technology Center, RWTH Aachen/Germany

The use of clinical decision support-systems (CDSS) raises ethical concerns already under significant debate. Many of the elements of these debates (such as unclear responsibility attribution, lack of explainability of decisions, and the impossibility of trusting machines) are discussed separately, without a more comprehensive perspective that brings them together. Autonomy is a useful tool for developing such a perspective. After all, autonomous agents are paradigmatically agents to whom responsibility is clearly attributable, and lack of explainability is clearly a concern when it prevents agents from thinking

for themselves. Thus, autonomy promises to provide a normative perspective from which to develop a comprehensive view on the use of CDSS. To develop this view, we first discuss the role of AI in medical decision-making and introduce the concepts of manipulation and autonomy support. We then argue that the use of CDSS can help physicians and patients think and decide for themselves and thereby be a form of autonomy support. It can do this principally by managing expert knowledge for making high-stakes decisions with limited resources, and by providing help for patients in understanding their situation. However, we also argue that CDSS uses can amount to undue manipulation and thereby threaten autonomy. They can do this by, e.g., limiting options in the name of easier decision-making, or providing biased, incomplete, or erroneous information to physicians and patients. We conclude that the use of CDSS is a double-edged sword that should be carried with caution. On the one hand, we should encourage the possibilities for better decisions and explore opportunities for AI to contribute to more robust clinical decision-making. On the other hand, this introduces the risk of the exact opposite: that we undermine autonomy through undue manipulation.

Reconfiguring Autonomy? Robotics And Care In A German Research Initiative

Maximilian Braun, Research associate and doctoral student, Munich School of Social Sciences and Technology, Technical University of Munich, Munich, Germany

Svenja Breuer, Research associate and doctoral candidate, Munich Center for Technology in Society, Technical University Munich, Munich, Germany

European and German funding discourses often frame the application of robots in healthcare environments as a means of enabling older adults to preserve a particular vision of a desirable life, one that follows an ethics of independence and self-determination. Here, the promise of a "relative autonomy" from "institutionalized care" for users of robots serves as a goal that innovators, policy-makers and funding programs strive to realize alike. However, the interplay between such an understanding of (human) autonomy and robotics-enhanced healthcare practices (and human care work more generally) seems not entirely clear: How might the widespread adoption of robots in healthcare settings re-conceptualize existing understandings of autonomy in human care? And what place do different stakeholders see for the envisioned technical autonomy of robotics systems in healthcare environments?

Looking at the case of a German research initiative for research on healthcare robotics, we explore a particular environment in which ideals of care-receivers' autonomy and self-determination dominate stakeholders' understandings and conceptualizations of care. Based on interviews with robotics researchers, project funders, caregiving experts and healthcare practitioners, we explore the different actors' understandings of autonomy, show how they are embedded (or not) in promissory narratives that support the initiative's research and how some of them rub up against the persisting material and social limitations of robots. Based on our findings, we point to the potential shortcomings that emerge from these understandings of autonomy and how they resonate with but also exceed the existing concerns around autonomy in technology policy.