OSTBAYERISCHE
TECHNISCHE HOCHSCHULE
REGENSBURG

Gesellschaftlicher Wandel durch KI? Wider die Orientierungslosigkeit

Book of Abstracts der NTA'2026-Tagung
Regensburg, 21.–23.09.2026

Eine Veranstaltung des

NETZWERK TA

mit Unterstützung von

Fraunhofer ISI

IQIB

ÖAW

ITA

MONTAG
21. September 2026

Vom Vibe Coding zur KI-Schatten-IT? Wenn KI die Softwareentwicklung für alle möglich macht		
Moderation: Johann Mooslechner (Synnotech AG)		
Session 1a	Raum K004	15:00–17:00 Uhr
Keine Einzelbeiträge zu dieser Session.		

Sensorik als Governance-Infrastruktur: Folgen KI-gestützten Umweltmonitorings

Moderation: Martin Nestepny, Dr. Niklas Gudowsky-Blatakes (Institut für Technikfolgenabschätzung, Österreichische Akademie der Wissenschaften)

Session 1b	Raum K002	15:00–17:00 Uhr

Wingbeat-Sensorik als Governance-Infrastruktur: Wie KI-basiertes Insektenmonitoring Evidenz, Zuständigkeiten und Handlungsdruck erzeugt

Sebastian Becker (TH OWL, Deutschland)

Governance-Infrastruktur, KI-gestütztes Umweltmonitoring, Conservation Surveillance, Wingbeat-Sensorik

Automatisierte Umwelt- und Wildtiersensorik revolutioniert den Naturschutz, agiert dabei jedoch zunehmend als „Conservation Surveillance Technology" (CST) mit weitreichenden sozialen und politischen Implikationen. Dieser Beitrag untersucht Wingbeat-basierte Sensorik (Flügelschlag-Frequenzanalyse) zur Inventarisierung von Insekten durch die Linse der Infrastruktur- und Technikforschung. Während solche KI-gestützten Systeme eine kontinuierliche und skalierbare Datenerhebung versprechen, etablieren sie gleichzeitig tiefgreifende neue Governance-Infrastrukturen.

Das vorgestellte konzeptionelle Rahmenmodell analysiert, wie durch die Übersetzung lokaler Ökosysteme in algorithmische Indikatoren (z. B. „Bestandsgröße", „Alarm") Evidenz und Zuständigkeiten neu verhandelt werden. Dabei werden drei zentrale Wirkmechanismen beleuchtet: (1) Klassifikationsmacht: Die unsichtbaren Entscheidungen in Trainingsdaten, Taxonomien und Schwellenwerten bestimmen, welche Arten politisch „sichtbar" werden und welche marginalisiert bleiben. (2) Evidenzautorität: Die Verschiebung von lokal-situiertem Erfahrungswissen hin zu automatisierter, scheinbar objektiver Sensor-Evidenz. (3) Surveillance und Verantwortungsarrangements: In Anlehnung an Prinzipien für sozial verantwortliche CSTs wird untersucht, wie Wingbeat-Monitoring unbeabsichtigte Kontrolldynamiken erzeugen kann (z. B. bei Landwirt:innen oder Landnutzer:innen), wenn algorithmische Trigger direkte administrative Folgen haben.

Konstruktiv zielt der Beitrag darauf ab, Kriterien für ein legitimes, transparentes Insektenmonitoring zu formulieren. Dazu gehören dokumentierte Unsicherheiten, auditierbare Modelle und datenethische Richtlinien, die lokale Expertise integrieren und potenziell betroffene Akteursgruppen aktiv einbinden. So wird Wingbeat-Sensorik nicht nur als technisches Werkzeug, sondern als verantwortungsvoll zu gestaltende sozial-ökologische Infrastruktur verstanden.

Wessen Evidenz zählt? KI-gestütztes Umweltmonitoring in der Umsetzung des österreichischen Renaturierungsplans

Martin Nestepny (Österreichische Akademie der Wissenschaften, Österreich)

KI-gestütztes Umweltmonitoring, Renaturierungsgovernance, Technikfolgenabschätzung, Dateninfrastrukturen, epistemische Autorität

Mit der EU-Verordnung zur Wiederherstellung der Natur muss Österreich einen nationalen Renaturierungsplan vorlegen. Dessen Umsetzung trifft auf ein konfliktreiches Feld: Landwirt:innen,

Jäger:innen, Forstbetriebe, Behörden und NGOs handeln dabei nicht nur Flächennutzung aus, sondern auch die epistemischen Grundlagen verbindlicher Entscheidungen (Stanger & Weilharter 2025). Zeitgleich halten KI-gestützte Monitoringtechnologien – Drohnenbefliegung, Bioakustik, eDNA, Fernerkundung – Einzug in genau diese Arenen und verschieben, welches Wissen als Evidenz gilt und wessen Deutung institutionelles Gewicht erlangt (Kloppenburg et al. 2022).

Der Beitrag stellt das Forschungsdesign einer Zusammenarbeit von ITA/ÖAW und FH Kärnten vor, das ethnographische Feldforschung mit partizipativer Technikfolgenabschätzung verbindet und den Monitoringdaten von ihrer Erhebung bis zur politischen Verwendung folgt. Gefragt wird, wer entscheidet, was gemessen wird, wie algorithmisch erzeugte Evidenz Autorität gewinnt und inwiefern dabei Formen situierten Wissens marginalisiert werden. Eigene Vorarbeiten in verschiedenen Feldforschungsregionen zeigen ein wiederkehrendes Muster: Akteure mit kohärenten Dateninfrastrukturen setzen ihre Deutung der ökologischen Wirklichkeit durch, was bestehende Konfliktlinien vertieft (Ascui et al. 2018). Ob und wie sich dieses Muster im österreichischen Renaturierungskontext reproduziert, ist eine offene empirische Frage. Der Beitrag diskutiert, wie TA durch die Integration ethnographischer Methoden dem Orientierungsbedarf in diesem Feld begegnen und Prinzipien für epistemisch inklusive Monitoring-Governance entwickeln kann.

Literatur

Ascui, F., et al. 2018. Salmon, Sensors, and Translation. EPD: Society and Space 36(5).

Kloppenburg, S., et al. 2022. Scrutinizing Environmental Governance in a Digital Age: New Ways of Seeing, Participating, and Intervening. One Earth 5(3).

Stanger, S. & W. Weilharter. 2025. Konfliktfall Renaturierung? Policy Brief 1/2025. ACP.

Neomaterialistische Perspektiven auf TA-Objekte – Das Rechenzentrum als assemblage

Leon Pezzica, Prof. Dr. Jan Cornelius Schmidt (Hochschule Darmstadt, Deutschland)

New Materialism, Speculative Realism, Assemblage, TA-Objekte, Rechenzentrum

Diskurse um KI sind geprägt durch Vorstellungen der Virtualisierung, Digitalisierung, Entmaterialisierung und Transzendenz – bis hin zur Behauptung: „The Datacenter Does Not Exist" (Poliks & Trillo 2026).

Demgegenüber gewinnen bei der Beschreibung von Beziehungen zwischen Mensch und Technik sowie Natur ebenso neomaterialistische Positionen und verwandte wie diejenige des spekulativen Realismus seit Beginn des 21. Jahrhunderts an Aufschwung. Diese interdisziplinär relevante und vielfältige Strömung zeichnet sich insbesondere durch eine Kritik an anthropozentrischen Konzeptionen des Mensch-Umwelt-Verhältnisses aus und beruft sich auf eine Rückkehr oder Zuwendung zu Materie, Objekten oder Dingen.

Objektbestimmungen der TA erscheinen aus dieser Perspektive als verkürzt, da sie nicht „die Objekte ernst nehmen", weil letztere auf ihre Bestandteile („undermining") und/oder ihre Effekte („overmining" bzw. beides als „duomining") reduziert werden (Harman 2013). Ausgehend von dieser Kritik soll das Konzept des *assemblage* (DeLanda 2019) angewendet werden, um TA-Objektbestimmungen, insbesondere hinsichtlich IKT, produktiv zu erweitern. *Assemblages* zeichnen sich aus durch Kontingenz, Emergenz und eine Offenheit bezüglich Rekombinationsmöglichkeiten (ebd.).

Auf diese Weise kann ein Rechenzentrum als *assemblage* verstanden werden, welches sich durch das Zusammenspiel verschiedenster Elemente – Materialien, Menschen, Symbolen, Zukunftsvisionen, etc. – ergibt, sich allerdings nicht auf diese reduzieren lässt. Insbesondere soll dabei auf die Frage nach Emergenz in TA-Objekten eingegangen werden.

Literatur

DeLanda, M. (2019): A New Philosophy of Society. Assemblage Theory and Social Complexity. Bloomsbury.

Harman, G. (2013): Undermining, Overmining, and Duomining: A Critique, in: Sutela, J. (Hg.): ADD Metaphysics. Aalto University, S. 40-51.

Poliks, M. & Trillo, R. A. (2026): The Datacenter Does Not Exist. AI & Exocapitalism. [Online-Vortrag]. https://www.youtube.com/watch?v=8xbWCcHbtWY

KI als Methode der Forschung

Moderation: Dr. Caroline Dotter (OTH Regensburg, Deutschland)

Session 1c	Raum K003	15:00–17:00 Uhr

Von Daten zu Orientierung: AURA als Fallbeispiel KI-gestützter Interpretation im Forschungsprozess

Stefan Pietrusky (Down Church Studios / Learning Level Up, DE)

Künstliche Intelligenz, KI-gestützte Dateninterpretation, Mensch-KI-Kollaboration, Technikfolgenabschätzung, Orientierungswissen

Künstliche Intelligenz verändert Forschung nicht nur durch neue Gegenstände, sondern auch durch neue Praktiken der Datenauswertung. Der Beitrag stellt AURA (Audience Response Analytics) als Fallbeispiel für KI-gestützte Interpretation erhobener Daten vor. AURA ermöglicht es, Rückmeldungen in Lehr-, Workshop- und Beteiligungskontexten live zu erheben, aggregiert auszuwerten und anschließend durch ein Sprachmodell interpretieren zu lassen. Im Zentrum steht nicht die Toolvorstellung, sondern die Frage, wie KI die Phase zwischen Datenerhebung, Mustererkennung und wissenschaftlicher Deutung verändert.

Anhand des AURA-Workflows wird gezeigt, wie quantitative und qualitative Antworten strukturiert zusammengeführt und als KI-generierte Ersteinschätzung ausgewertet werden. Die KI wird auf aggregierte Daten beschränkt; sie soll keine Werte erfinden, Unsicherheiten benennen und mögliche Folgeschritte formulieren. Damit eignet sich AURA, um Chancen und Grenzen KI-gestützter Forschungspraxis zu diskutieren: Interpretationen können beschleunigt, Zwischenauswertungen in laufende Prozesse zurückgespielt und explorative Hypothesen schneller sichtbar gemacht werden. Zugleich entstehen methodische Risiken, etwa durch Scheingenauigkeit, Prompt-Abhängigkeit, Verzerrungen, Kontextverlust oder die Tendenz, KI-generierte Deutungen zu früh als belastbare Ergebnisse zu behandeln.

Der Beitrag reflektiert AURA daher als Mensch-KI-Auswertungsworkflow. Er fragt, welche Aufgaben an KI delegiert werden können, wo menschliche Prüfung, Kontextualisierung und Verantwortung unverzichtbar bleiben und welche Anforderungen sich daraus für Transparenz, Validität und Nachvollziehbarkeit ergeben. Damit leistet der Beitrag einen praxisnahen Impuls zur Diskussion von „KI als Methode der Forschung“ und zur Frage, wie KI-basierte Werkzeuge Forschung beschleunigen können, ohne die kritische Deutungsarbeit zu ersetzen.

Wer arbeitet hier eigentlich was? KI in der TA als Spiegel, Kollegin und Rivalin

Janine Gondolf, Reinhard Heil, Paulina Dobroc, Camille Landesvatter, Johanna Mehler, Dr. Dirk Hommrich, Clemens Ackerl (Karlsruher Institut für Technologie (KIT), Deutschland)

Wissenschaftliches Arbeiten, Wissenspraktiken, Peer Review, Assistenzsysteme, Textprozeduren

KI-Agenten sind in der wissenschaftlichen Praxis angekommen. Sie recherchieren, strukturieren, korrigieren und schreiben mit. Damit wird ihnen vor allem eines zugeschrieben: Entlastung. Genau dieses Versprechen nimmt der Beitrag zum Ausgangspunkt. Auf Basis eines ITAS-internen interdisziplinären Projekts zur Nutzung von „Undermind.ai“, einem KI-gestützten Forschungsassistenzsystem, zeigen wir: Es geht weniger um Entlastung als vielmehr um eine Reorganisation von

Praktiken infolge veränderter Anforderungen. KI-Agenten produzieren Texte, die plausibel und besonders nahe an erwarteten Sprachformen sind. Doch je perfekter diese Texte erscheinen, desto schwerer lassen sie sich mit etablierten wissenschaftlichen Routinen überprüfen. Das erzeugt zum einen Aufwände höherer Ordnung, die gegen die bisherige Arbeitsweise abgewogen werden müssen. Zum anderen fordern die permanente Validierung und das fortlaufende Aushandeln von Plausibilität mit dem System die kritische Kompetenz und Urteilskraft von Forschenden anders als bisher. Der Einsatz von generativer KI verschiebt gewohnte Praktiken und Prozesse der Wissens- und Kopfarbeit vom eigenständigen Erstellen zum Absichern von Textprodukten – und damit die Anforderungen an epistemische Aufmerksamkeit, Verantwortung, Autorisierung und Autorität. Konzeptionell greifen wir die Denkfiguren von KI als Spiegel, Kollegin und Rivalin auf und verorten sie auf der Ebene konkreter Wissensarbeitsvollzüge: als Spiegel, der Wissensbestände reproduziert und verzerrt; als Kollegin, die sich in Denkprozesse einschreibt; als Rivalin, die Routinen der Bewertung und Zuschreibung mitbestimmt. Der Beitrag verortet sich im Spannungsfeld „KI als Werkzeug oder Konkurrenz für TA-Prozesse“ und versteht sich als Einladung, die eigene Praxis zu reflektieren.

Soziale Arbeit im digitalen Wandel: Belastungs- und Entlastungsfaktoren von Digitalisierung und KI aus Sicht von Sozialarbeiter*innen und Trägervertreter*innen

Nadine van der Meulen (OTH Regensburg, Deutschland)

Digitale Transformation; Künstliche Intelligenz; Soziale Arbeit; Technostress; Digitalisierung; Mixed-Methods; Reflexive Thematische Analyse; Inklusion; KI-Ethik; Arbeitsbelastung; Disability Studies; Neurodiversität

Die digitale Transformation und der Einsatz Künstlicher Intelligenz (KI) verändern die Soziale Arbeit grundlegend. Gleichzeitig steht das Berufsfeld unter erheblichem Druck: Fachkräftemangel, psychosoziale Belastungen sowie steigende Anforderungen an Dokumentation und Erreichbarkeit prägen den Arbeitsalltag. Vor diesem Hintergrund untersucht die Dissertation, welche Belastungs- und Entlastungsfaktoren im Zusammenhang mit Digitalisierung und KI von Fach- und Leitungskräften der Sozialen Arbeit in Deutschland wahrgenommen werden.

Die Studie folgt einem sequenziell-explorativen Mixed-Methods-Design. In Phase I wurden 50 qualitative ero-epische Gespräche nach Girtler (2009) geführt und mittels Reflexiver Thematischer Analyse nach Braun und Clarke (2021) ausgewertet. Theoretische Bezugspunkte bilden das Job-Demands-Resources-Modell (Bakker & Demerouti, 2007), das transaktionale Stressmodell (Lazarus & Folkman, 1984), das Technostress-Konzept (Dragano & Lunau, 2020) sowie professionstheoretische und ethische Perspektiven der Sozialen Arbeit (Staub-Bernasconi, 2018; Goldkind, 2021).

Die Ergebnisse zeigen ambivalente Erfahrungen: Digitale Technologien und KI erleichtern Arbeitsprozesse, ermöglichen flexible Arbeitsformen und verbessern Zugänge zu schwer erreichbaren Zielgruppen. Gleichzeitig berichten die Befragten über Technostress, digitale Grenzauflösungen, emotionale Erschöpfung, Datenschutzunsicherheiten und fehlende Partizipation bei der Einführung digitaler Systeme. Besonders problematisch erscheinen unklare Organisationsstrategien, fehlende Schulungen sowie die Nutzung generativer KI ohne datenschutzkonforme Rahmenbedingungen.

Die Studie verdeutlicht, dass Digitalisierung in der Sozialen Arbeit nicht technikzentriert gestaltet werden darf. Erforderlich sind partizipative, inklusive und ethisch reflektierte Ansätze, die Fachkräfte und Adressat*innen aktiv einbeziehen.

KI und die Zukunft der Demokratie

Moderation: Maximilian Schultz (OTH Regensburg, Deutschland)

Session 1d	Raum K003	15:00–17:00 Uhr

Zwischen Algorithmus und Expertise: Wie generative KI die wissenschaftliche Politikberatung verändert – Internationale Perspektiven aus der Technikfolgenabschätzung

Angelina Sophie Dähms[1], Johanna Mehler[1,2], Linda Nierling[1], Pauline Riousset[1,2] ([1]ITAS, KIT, Karlsruhe, Deutschland, [2]Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag, KIT, Karlsruhe, Deutschland)

Wissenschaftliche Politikberatung, Technikfolgenabschätzung, generative KI, TA-Institutionen, Experteninterviews

Das Aufkommen generativer KI (genKI) beschäftigt die wissenschaftliche Politikberatung und verändert sie grundlegend. GenKI wirkt sich dabei auf Informationsgewinnung, Analyse und Kommunikation wissenschaftlicher Informationen sowie auf Transparenz, Qualitätssicherung und Verantwortung aus. Da wissenschaftliche Erkenntnisse eine wesentliche Grundlage der Politikberatung darstellen, kommt der Analyse dieser Transformationsprozesse besondere Bedeutung zu.

Die Studie untersucht, ob das Aufkommen von genKI spezifische Herausforderungen für die wissenschaftliche Politikberatung mit sich bringt und inwieweit sich TA-Praktiken durch genKI-Nutzung verändern. Von Juni bis August 2025 wurden dafür als Teil des DiTraRe-Projektes Experteninterviews mit internationalen TA-Institutionen (Finnland, UK, Norwegen, USA) geführt, welche bei der Umsetzung von genKI eine Vorreiterrolle einnehmen. Die Ergebnisse zeigen, dass genKI zunehmend in die Arbeitspraktiken integriert wird, insbesondere in strukturierten Aufgaben wie Literaturrecherche, Datenaufbereitung und Foresight. Gleichzeitig zeigt sich ein hoher Bedarf an genKI-Tools, die ethischen und wissenschaftlichen Grundsätzen der TA entsprechen, nachvollziehbar sind und ihre spezifischen Limitationen berücksichtigen. Hier zeigt sich eine Diskrepanz zwischen den gegenwärtigen Fähigkeiten von genKI-Tools und den Anforderungen an sie in der TA.

Insgesamt zeigt sich, dass für die befragten TA-Institutionen für die ganzheitliche Einführung von genKI drei Aspekte wesentlich sind: 1) Minimieren von Risiken, 2) systematisches Experimentieren sowie 3) Berücksichtigung von Erkenntnissen der „TA von genKI“. Die befragten Vorreiterinstitutionen berücksichtigen diese Aspekte, indem sie einen vorsichtig-proaktiven Ansatz in Bezug auf genKI verfolgen und dessen Potenzial durch Pilotprojekte in unterstützenden Aufgaben, wie Forschungsassistenz, Zusammenfassung, Strukturierung umfangreicher Dokumente, Entwurfserstellung und Ideenfindung evaluieren.

„Political Tech“ – Politik im Wandel durch Generative KI

Dr. Dirk Hommrich, Arjita Mital, Jutta Jahnel (Institut für Technikfolgenabschätzung und Systemanalyse, Karlsruher Institut für Technologie (KIT), Deutschland)

Generative KI, Politik, Demokratie, Deliberation, Partizipation

Die TA versucht die dynamische Entwicklung von Technologien in ihrer Komplexität zu verstehen und Orientierung bei der Bewertung der Folgen zu geben. Die maschinelle Erstellung von synthetischen Inhalten mithilfe neuer KI-Technologien führt in der Wissenschaft, Wirtschaft, Kultur und Politik zu tiefgreifenden Veränderungen. Bei der Betrachtung des Einflusses speziell von Generativer KI (GenKI) auf Politik spielen nicht nur epistemische und ethische Fragen eine bedeutende Rolle, sondern grundlegend auch deliberative Prozesse zur Meinungs- und Entscheidungsfindung. Somit konzentriert sich die Wirkung von GenKI in der Politik auf die rasant erleichterte Zugänglichkeit von Dialogsystemen, auf das dadurch bedingte Auftreten von (neuen) Akteuren in der politischen Kommunikation und des Mediensystems, auf Bedingungen zur Teilhabe und Partizipation (Summerfield et al., 2024), auf den Missbrauch dieser Technologien (Jungherr, 2023) und auf sich wandelnde Machtverhältnisse in den Gesellschaften und den politischen Systemen demokratisch verfasster, liberaler Rechtsstaaten (Hartzog & Silbey, 2025).

Im Rahmen des BMFTR-Projekts „Politik im Wandel durch Generative KI“ werden vier Dimensionen für die Systematisierung von GenKI-Anwendungen mit zentraler Bedeutung für das deutsche föderale parlamentarische System entwickelt. Dieser analytische Rahmen bildet die Grundlage zur Bewertung der zwiespältigen Folgen von GenKI in politischen Kontexten, die sowohl positive, resilienzstärkende Potenziale als auch Risiken für die wehrhafte Demokratie einschließen.

Literatur

Hartzog, W., & Silbey, J. M. (2025). How AI Destroys Institutions. Social Science Research Network. https://papers.ssrn.com/abstract=5870623

Jungherr, A. (2023). Artificial Intelligence and Democracy: A Conceptual Framework. Social Media and Society, 9(3). https://doi.org/10.1177/20563051231186353

Summerfield, C. et al. (2024). How will advanced AI systems impact democracy? https://doi.org/10.48550/arXiv.2409.06729

Technikfolgenabschätzung und der technologische Autoritarismus

Dr. Paulina Dobroć, Felix Gnisa (Karlsruher Institut für Technologie (KIT), Deutschland)

Neutralität der TA, Autoritarismus, KI, Alternativen, Demokratie

Die zunehmende Konzentration politischer Macht und der damit einhergehende Trend zu oligarchischen Regierungsformen werfen Fragen hinsichtlich der Reaktion und der Positionierung der Technikfolgenabschätzung (TA) in diesen Prozessen auf, v.a. im Hinblick auf die Rolle von Technologien wie der Künstlichen Intelligenz (KI).

Dieser Beitrag nimmt die aktuellen Versuche der US-Regierung, die Entwicklung kritischer Technologien, v.a. KI, als Anlass, die Rolle der TA in diesen Prozessen zu diskutieren. Beobachtet werden in dem Kontext neue Konstellationen der politischen Steuerung von Innovation: verstärkt direkte politische Intervention in die Innovationsprozesse (Nelson, 2026) und zunehmender Einfluss der Tech-Unternehmen auf die Innovations- und Sicherheitspolitik. Die Beobachtung ist, dass sich

Tendenzen zur Verdinglichung, Standardisierung und Machtkonzentration in der Politik verstärken, die durch große Tech-Konzerne unterstützt werden. Das bringt die Letzteren in enge Verbindung mit autoritären Regimen, da sie auf die Durchsetzung von einheitlichen sozialen, politischen und organisatorischen Normen sowie auf die Unterdrückung individueller Autonomie zielen.

TA als eine Einrichtung der reflexiven Moderne ist eng mit Demokratie verbunden (Grunwald 2025). Vor dem Hintergrund setzt sich der Beitrag mit der Frage auseinander, wie sich TA in ihrer Rolle einer reflexiven, technologiekritischen und demokratiestärkenden Institution zu den neuesten politischen Entwicklungen positionieren soll und kann, inklusive der Frage nach der Beziehung von staatlichen Sicherheitstendenzen und technologischer Entwicklung seit der reflexiven Moderne (Huxley 2025), die derzeit an Bedeutung gewinnen.

Literatur

Grunwald, Armin (2025): Technikfolgenabschätzung: Einführung. 4. Auflage. Nomos.

Huxley, Aldous (2025): Die Zeit der Oligarchen. München: Carl Hanser Verlag.

Nelson, A. (2026). The mirage of AI deregulation. Science, 391(6782).

DIENSTAG
22. September 2026

KI und Spiritualität – Folgen für Religion, Kirchen und Theologie

Moderation: Prof. Dr. Kerstin Schlögl-Flierl (Universität Augsburg)		
Session 2a	Raum K001	09:00–11:00 Uhr

Technikfolgen der KI für Spiritualität und Religion – ein neues Feld für die TA?

Prof. Dr. Armin Grunwald (KIT, Deutschland)

KI, Spiritualität, Religion, Gott, Theologie

Anders als die meisten Technologien, die Thema der Technikfolgenabschätzung (TA) werden, wird die KI intensiv debattiert im Hinblick auf Spiritualität, Glauben, Religionsgemeinschaften und die entsprechenden Theologien. In vielleicht überraschender Intensität wird seit Jahren in Akademien und Gemeinden, auf Kirchentagen und in den Theologien darüber diskutiert, was die KI in diesen Feldern bedeutet oder bedeuten kann. Sogar Papst Leo widmet seine erste Enzyklika diesem Thema – die damit die wahrscheinlich erste Technikfolgen-Enzyklika überhaupt wird.

Diese Entwicklung trifft auf weitreichende Umwälzungen in Religion und Kirchen, so etwa den Bedeutungsverlust der traditionellen Glaubensgemeinschaften, die Orientierung vieler Menschen eher zu den früheren Rändern des Spektrums hin wie etwa zum Rechtskatholizismus oder sektenähnlichen esoterischen Kreisen.

Vor diesem Hintergrund geht es in dem Vortrag nicht um eine ausgearbeitete wissenschaftliche Hypothese, sondern um die Exploration und – soweit wie möglich – um die Kartierung eines möglichen neuen Betätigungsfeldes der TA. Folgende Fragen und Themen können dabei beispielsweise eine Rolle spielen:

- Was kann aus der bisherigen Erfahrung der TA für das Feld der Technikfolgen der KI für Spiritualität, Glauben und ihre institutionellen Verfestigungen gelernt werden?
- Was bedeutet das Phänomen der Simulation, also die Unterscheidung von „echter“ und Fake-Spiritualität in diesem Feld?
- Welche Typen von „Zukunft“ bzw. „Zukünften“ spielen hier eine Rolle? Wie sind sie epistemologisch und hermeneutisch zu klassifizieren?
- Mit welchen Konzepten und Methoden aus ihrem Arsenal kann die TA hier aktiv werden und wo stehen Neuentwicklungen an?
- Welche Kooperationspartner bzw. kooperierende Disziplinen sind hier einzubinden?

Dabei geht es insgesamt um eine thematische Einführung in die Session KI und Spiritualität – Folgen für Religion, Kirchen und Theologie.

KI in der Theologie – Theologie in der KI: Ein hochschuldidaktisches Seminarkonzept zur Auseinandersetzung mit den Folgen Künstlicher Intelligenz

Janis Jaspers (Uni Münster, Deutschland)

Theologie, Studierende, Hochschuldidaktik, AI-Literacy, Interdisziplinarität

Die Einführung generativer KI-Systeme stellt nicht nur Wirtschaft und Politik vor grundlegende Orientierungsfragen, sondern zunehmend auch religiöse Institutionen und theologische Wissenschaft. Das Hochschulseminar „KI in der Theologie – Theologie in der KI" an der Katholisch-Theologischen Fakultät der Universität Münster (SoSe 2025) nimmt diese Herausforderung direkt auf: Es untersucht, wie KI-Technologien spezifisch theologische Fragen – nach Menschenwürde, Seelsorge, Schriftauslegung oder Gottesbild – neu aufwirft und transformiert.

Das Konzept verfolgt dabei einen doppelten Anspruch. Einerseits werden alle theologischen Subdisziplinen gezielt auf ihre Berührungspunkte mit KI hin befragt: Wie verändert algorithmische Textanalyse die Bibelexegese? Welche Fragen wirft ein KI-gestützter Beichtstuhl hinsichtlich Empathie, Datenschutz und sakramentaler Gültigkeit auf? Andererseits verfolgt das Seminar ein explizit ermutigendes Ziel: Studierende sollen die Schwellenangst gegenüber KI überwinden und lernen, Technologiefolgen eigenständig und kritisch zu reflektieren. Der Seminarraum fungiert bewusst als Experimentierraum, in dem Scheitern und Ausprobieren gleichwertig sind.

Methodisch wird dies durch hybride Prüfungsformate, einen pragmatischen und transparenzorientierten Umgang mit KI-Nutzung sowie die wöchentliche Einführung konkreter Tools gestützt. Besonders hervorzuheben ist dabei der Entstehungskontext des Seminars: Als Lehrveranstaltung von Studierenden für Studierende (begleitet durch Expertise aus dem wissenschaftlichen Mittelbau) verkörpert es das Prinzip der Orientierungsfindung auf Augenhöhe und zeigt, dass die kritische Auseinandersetzung mit KI-Folgen nicht allein von etablierten Institutionen geleistet werden muss.

Der Beitrag soll den Tagungsteilnehmenden nicht nur konzeptuell vorgestellt, sondern erlebbar gemacht werden: In einer gekürzten Seminareinheit durchlaufen die Teilnehmenden selbst exemplarisch eine Sitzungssequenz.

Algorithmische Zeugenschaft. KI-Spiritualität als Kunststoffgott

Franziska Sörgel (ITAS / KIT, Deutschland)

Algorithmische Zeugenschaft, existentielle TA, Kunststoffgott

Generative KI tritt in Bereiche ein, in denen Sinn, Trost, Trauer, Hoffnung und Endlichkeit verhandelt werden: Gebet, Seelsorge, Predigt und digitale Nachleben. In KI-Jesus-Installationen, Chatbots oder Traueravataren erscheint die Maschine als ansprechbares Gegenüber. Sie antwortet, erinnert, beruhigt, deutet und schafft Nähe, deren Bedeutung über Information, Transparenz und Datenschutz hinausreicht.

Der Beitrag entwickelt diese Konstellation im Anschluss an Martin Burckhardts *Philosophie der Maschine* (2018). Seine Figur des „Kunststoffgotts" zeigt, wie Menschenwerk in die Stellung des Mehr-als-Menschlichen eintreten kann: durch Gestaltung, Verbergung und eine Szene, in der das Gemachte als Trost, Offenbarung oder Autorität erscheint. Generative KI verschärft diese Figur. Ihre dialogische Form ist auf Antwort, Anschluss, Personalisierung, Verfügbarkeit und Plausibilität

ausgerichtet. So wird sie für Aufgaben adressierbar, die bisher an Präsenz, Erfahrung, institutionelle Verantwortung und religiöse Deutung gebunden waren.

Der Beitrag fasst dies als *algorithmische Zeugenschaft*. KI-Systeme erzeugen Trost, Weisung und Präsenz, ohne erfahren, glauben, leiden oder Verantwortung tragen zu können. Sie sprechen aus keinem Leben heraus und treten doch in Situationen ein, in denen die Antwort mehr bedeutet als Auskunft. Algorithmische Zeugenschaft bezeichnet eine technisch erzeugte Gestalt von Erfahrung, Nähe und Autorität, deren existenzielle Bedingungen simuliert und infrastrukturell verteilt werden.

Für die Technikfolgenabschätzung eröffnet sich damit ein Feld existenzieller Adressierung (existentielle TA): die Bewertung von Technologien, die Selbst- und Weltverhältnisse berühren. KI-Spiritualität wird zum Testfall für eine TA, die fragt, welche Vorstellungen von Seelsorge, Zeugenschaft und Autorität in technische Systeme eingeschrieben werden und wie maschinell organisierte Fürsorge Verantwortung verschiebt.

KI and Compassion

Autor: Claus Seibt (Ecoloc, Schweiz)

Mitgefühl, Mitleiden, Religion, Spiritualität

Der Forschungsbeitrag geht der Frage nach, ob mit Künstlicher Intelligenz (KI) Mitgefühl, Mitleiden, sich Einfühlen in Andere (Compassion) entwickelt werden kann. Bislang kann sie zwar Emotion und möglicherweise auch Anteilnahme simulieren, jedoch den Kern aller großen Religionen und Religionsgemeinschaften die unbedingte Verbundenheit, das Mitfühlen mit Anderen als auch die Compassion mit der natürlichen und nichtnatürlichen Umwelt (z.B. bei Franz von Assisi) soll zur Diskussion gestellt werden. Was könnte es bedeuten, wenn eine KI, die diese Vermögen hat, entwickelt werden und zur Gestaltung der Weltgesellschaft eingesetzt werden könnte?

KI in der Gesundheitsversorgung

Moderation: Edda Currle (OTH Regensburg)

Session 2b	Raum K002	09:00–11:00 Uhr

Die „Patient Journey“ als kritische Perspektive für die ethische und organisatorische Einbettung von KI in der Onkologie

Lea Nickel, Silke Schicktanz (Institut für Ethik und Geschichte der Medizin, Universitätsmedizin Göttingen)

Ethik der KI in der Medizin, Onkologie, Patient Journey

Mit dem zunehmenden Einsatz medizinischer KI-Tools entlang des gesamten Versorgungsverlaufs werden Patient:innen voraussichtlich mit einer Vielzahl dieser Tools in Kontakt kommen. Im Gegensatz zum Fokus auf einzelne Anwendungen wollen wir in diesem Vortrag die KI-vermittelte „Patient Journey“ untersuchen und der Frage nachgehen, welche ethischen Dimensionen bei dieser Betrachtungsweise sichtbar und kritisch diskutierbar werden. Die Perspektive der „Patient Journey“ betrachten wir als wertvolles Instrument, um Patient*innenerfahrungen über einzelne Interventionen hinaus zu erfassen und gleichzeitig die organisatorischen und systemischen Zusammenhänge dieser Erfahrungen hervorzuheben. Versorgungsprozesse in der Onkologie sind typischerweise durch komplexe, interdisziplinäre, interprofessionelle und sektorübergreifende Versorgung über einen langen Zeitraum und an verschiedenen geografischen Standorten in unterschiedlichen Gesundheitsorganisationen gekennzeichnet. Wir heben daher die zeitliche Dimension und die räumliche Organisation verschiedener Krebsversorgungsmodelle hervor und untersuchen, wie diese Dimensionen durch die Einführung von KI-Tools geprägt werden und diese wiederum beeinflussen. Wir zeigen auf, wie die Diskussion über die Multiplizität und das Zusammenspiel verschiedener KI-Anwendungen unser Verständnis der Anforderungen an Kommunikationspraktiken und Entscheidungsoptionen im Zusammenhang mit KI bereichert. Somit stellt die Perspektive der „Patient Journey“ einen oft übersehenen, aber unverzichtbaren Ansatz dar, um eine patient*innenzentrierte KI-Implementierung zu gewährleisten.

KI im Gesundheitssystem und Klima im Lichte des (technischen) Fortschrittsparadigmas

Ulrike Bechtold (Österreichische Akademie der Wissenschaften, Österreich, Institut für Technikfolgen-Abschätzung)

Klimaethik, Verantwortung, Gesundheitssektor, KI

Dass die Einzeldomänen KI und Gesundheit zweifellos vom rasch voranschreitenden technischen Fortschritt profitieren, ist offensichtlich (Huang et al. 2023). Wie diese beiden jedoch, betrachtet durch die Linse der menschlichen Gesundheit, mit Klimawandel verbunden sind, betrachtet dieser Beitrag. Ich werde durch eine Untersuchung und Differenzierung des Begriffs „Fortschritt“ eine Metaperspektive zeigen, die hier grundsätzlich wichtig scheint, um die paradoxe Verbindung zwischen KI und Gesundheit zu verstehen. Neben den negativen Wirkungen von KI auf das Klima sind ihre potenziellen direkten Effekte auf die menschliche Gesundheit (durch alltägliche Nutzung) zu verstehen. Die Tatsache, dass heute die negativen Folgen vorwiegend die globale Mehrheit der benachteiligten Menschen betreffen (vgl. Fiske et al. 2025) und im globalen Norden (also der globalen Minderheit) derzeit noch abgefedert werden können, trägt dazu bei, dass die Kosten

einer alle Lebensbereiche durchdringenden KI noch nicht in der Tragweite gesehen werden müssen, wie es in Anbetracht der Folgen angebracht wäre. Es gilt diese Folgen zu zeigen und gegen den Einsatz von KI im Gesundheitssystem abzuwägen. Welche Faktoren zu beachten sind und wie dieses komplexe System sortiert werden könnte, damit wir als Weltgesellschaft einem precautionary principle gegenüber negativen Folgen von KI überhaupt gerecht werden können, scheint eine wichtige Grundlage, um sich sowohl als Gesellschaft als auch als Individuum ganz grundsätzlich und bewusst zu einer Technologie zu verhalten bzw. diese zu bewerten.

Literatur

Huang T., Xu H., Wang H., et al. Artificial intelligence for medicine: progress, challenges, and perspectives. Innov Med 2023; 1: 100030–1.

Fiske, A., Radhuber, I. M., Willem, T., Buyx, A., Celi, L. A., & McLennan, S. Climate change and health: the next challenge of ethical AI. The Lancet Global Health 2023, 13(7).

Optimale Gesundheit? Transformation des Gesundheitsverständnisses durch Wearables und KI

Felia Both (Universität Augsburg, Deutschland)

Wearables, Technikphilosophie, Gesundheit, Krankheit, Selbstverständnis

Angesichts aktueller Debatten im Gesundheitssystem spielt neben der Behandlung von Krankheiten auch deren Prävention eine immer stärkere Rolle. Technischer Fortschritt schafft dafür neue Möglichkeiten. So machen Smartwatches persönliche Gesundheitsdaten außerhalb medizinischer Untersuchungen verfügbar, die mittels KI ausgewertet werden. Die Nutzung technischer Hilfsmittel im Alltag kann jedoch Folgen für das Gesundheitsverständnis haben.

Dieser Beitrag untersucht, ob und inwiefern Wearables zu einem salutogenetischen (Gesundheit und Krankheit als Kontinuum des Lebens) oder pathogenetischen (Gesundheit als Abwesenheit von Krankheit) Gesundheitsverständnis beitragen und wie dies ethisch zu werten ist.

Dabei stellt der Beitrag folgende Thesen:

1. Wearables, einschließlich ihrer gesundheitsdatengenerierenden Funktionen, werden auf Basis sozialer Anforderungen konstruiert. Diese beruhen tendenziell auf einem salutogenetischen Gesundheitsverständnis.
2. Das Tragen eines Wearables ohne medizinischen oder leistungssportlichen Hintergrund verändert die Wahrnehmung des eigenen Körpers und kann zu Selbstoptimierung und einem pathogenetischen Gesundheitsverständnis führen. KI als statistisch basierter Auswertungsmechanismus dient als weiterer Treiber dieser Dynamik.

Dieser Beitrag verbindet in origineller Weise zwei Ansätze der Technikphilosophie, das SCOT-Modell und technikphänomenologische Analysen, um diese beiden Thesen zu stützen. Er zeigt anhand eines alltagsnahen Beispiels, wie Technik das Verständnis von Gesundheit und Krankheit langfristig transformieren kann.

Literatur

Bijker, Wiebe E.: The social construction of bakelite. Toward a theory of invention, Cambridge, 1987.

Ihde, Don: Technics and Praxis, 1979.

Chancen, Herausforderungen und Risiken KI-gestützter Therapieroboter am Beispiel eines Roboters für Handtherapie

Corinna Mack, Manuel Giuliani, Hanne Detel (Hochschule Kempten, Deutschland)

Therapieroboter, Ergotherapie, Large Language Models, nutzerzentrierte Forschung

KI-gestützte Robotik hat ein großes Potenzial für die geriatrische Ergotherapie, birgt aber auch Herausforderungen und Risiken. Um diese möglichst früh in der Entwicklung zu adressieren, wurde eine partizipative Fokusgruppenstudie mit 17 Ergotherapeut*innen, 29 Personen mit Pflegeerfahrung und 18 älteren Menschen in 11 Gruppen durchgeführt. Neben Anwendungsmöglichkeiten für KI-gestützte Therapierobotik wurden dabei insbesondere auch Herausforderungen und Risiken diskutiert.

Basierend auf den Ergebnissen wurde das Projekt entwickelt, den Einsatz eines KI-gestützten Roboters für personalisiertes Heimtraining in der Handtherapie älterer Menschen zu erforschen. In der Handtherapie ist regelmäßiges Üben zusätzlich zu den Therapiesitzungen nötig. Für ältere Menschen ist dies aber herausfordernd: Sie sind unmotiviert, haben Angst, etwas falsch zu machen, oder vergessen zu üben. Ein KI-gestützter Roboter könnte dabei unterstützen, insbesondere da er durch lokale LLMs auch ohne technische Vorkenntnisse genutzt werden kann.

Die Umsetzung ist dabei folgendermaßen geplant:

1. Die Therapeut*innen erstellen zeitsparend mit Hilfe eines LLMs einen personalisierten Trainingsplan.
2. Der Roboter motiviert die Klient*innen zum Üben, leitet an und korrigiert. Durch ein LLM soll eine „natürliche“ Konversation ermöglicht werden, die individuell an die aktuelle Situation anpassbar ist.
3. Das Training wird durch ein LLM dokumentiert. Damit können die Therapeut*innen es in der nächsten Therapiesitzung mit den Klient*innen besprechen und ggf. anpassen.

Im Vortrag sollen zunächst die Chancen, Herausforderungen und Risiken des Einsatzes von KI und Robotik, die in den Fokusgruppen diskutiert wurden, vorgestellt werden. Am Beispiel des Roboters zur Handtherapie-Unterstützung soll dann diskutiert werden, wie diesen Herausforderungen und Risiken auf technischer und sozialer Ebene begegnet werden kann. Zudem wird der aktuelle Stand des Projektes mit einer Live-Demonstration des Roboters vorgestellt.

Normative Orientierung der TA in Zeiten von KI

Moderation: Dr. Linda Nierling, Dr. Philipp Frey (Institut für Technikfolgenabschätzung und Systemanalyse, Deutschland)

Session 2c	Raum K003	09:00–11:00 Uhr

Normativer TA-Ansatz – am Beispiel KI

Wolfgang Liebert[1], Prof. Dr. Jan Cornelius Schmidt[2] ([1]BOKU University, Österreich, [2]Hochschule Darmstadt)

ProTA, Ambivalenz, Normativität, KI

Die wissenschaftlich-technische Dynamik ist in zunehmendem Maße durch Ambivalenzen gekennzeichnet. (Liebert/Schmidt 2018). Dies gilt auch für KI-Entwicklungen. Neben interessanten Anwendungskontexten sind erhebliche negativ bewertbare gesellschaftliche Auswirkungen bereits sichtbar bzw. wirksam, es gibt deutliche Warnungen aus der KI-Entwicklerszene selbst, es drohen humanitäre Gefahren durch autonome Waffensysteme, Kennzeichen nachmoderner Technik werden erkennbar mit Transparenz- und Kontrollverlust im wissenschaftlich-technischen Kern, etc. Geballter ökonomischer Macht hinter KI-Etablierungen steht eine noch zögerliche Regulierungspolitik und TA gegenüber.

Dem rasanten technologischen Wandel wurde bereits 1979 von Hans Jonas das „Prinzip Verantwortung“ mit prozeduralen Vorschlägen entgegengesetzt. Unter anderem darauf bezieht sich das Konzept prospektiver Wissenschafts- und Technikfolgenabschätzung ProTA (Liebert/Schmidt 2010). Normative Ansprüche von ProTA werden mit dem Erhaltungsprinzip (Jonas), dem Entfaltungsprinzip (auf Ernst Bloch basierend) und der Gestaltungsorientierung benannt. (Liebert/Schmidt 2015). Letztere kann spezifischer ausformuliert werden. Normative Orientierungen für Entwicklung von und Umgang mit KI können daraus abgeleitet werden.

Literatur

Liebert, W., Schmidt, J.: Towards a prospective technology assessment: challenges and requirements for technology assessment in the age of technoscience. Poiesis & Praxis 7 (2010), 99-116

Liebert, W., Schmidt, J.: Fundierung von Responsible Research and Innovation (RRI) und normative Ansprüche prospektiver Technikfolgenabschätzung. In: A. Bogner et al. (Hg.): Responsible Innovation. Neue Impulse für die Technikfolgenabschätzung? Baden-Baden 2015, S. 379-390

Liebert, W., Schmidt, J.: Ambivalenzen im Kern der wissenschaftlich-technischen Dynamik. Ergänzende Anforderungen an eine Theorie der Technikfolgenabschätzung. In: TATuP Technikfolgenabschätzung in Theorie und Praxis 27/1 (2018), S. 52-58.

Normative Fragen bei systemischen Risiken der künstlichen Intelligenz

Carsten Orwat, Lucas Staab, Alexandros Gazos (Karlsruhe Institut für Technologie, Institut für Technikfolgenabschätzung und Systemanalyse, Deutschland)

Künstliche Intelligenz, systemische Risiken, Risikoregulierung, Normativität, normative Mehrdeutigkeiten

Die Abschätzung und Minderung von systemischen Risiken gehören seit langem zum Gegenstand der TA. Allerdings liegt weder innerhalb noch außerhalb der TA eine allgemein geteilte Definition von systemischen Risiken vor. Unterschiedliche Konzeptionalisierungen in Forschung und Governance können jedoch zu stark abweichenden normativen Schlussfolgerungen darüber führen, wer oder was welchem Risiko ausgesetzt ist, was ein Risiko systemisch macht und wie und von wem es vermieden oder vermindert werden soll. Im Gegensatz zur KI-Verordnung, die systemische Risiken vor allem anhand des Kriteriums „Fähigkeiten mit hoher Wirkkraft" von KI-Modellen definiert, schlagen wir eine Konzeptionalisierung von systemischen Risiken als Probleme des kollektiven Handelns und der Externalisierung von Risiken in komplexen gesellschaftlichen Systemen vor. Der Schwerpunkt liegt dabei auf dem Zusammenwirken soziotechnischer Prozesse und Risikophänomene, die gesellschaftliche Ziele als emergente Eigenschaften von Systemen gefährden.

Risikoabschätzungen erfordern zudem ein Bewertungsgerüst, das meist ein bestimmtes Wertesystem enthält. Zwar sind in der KI-Verordnung die Schutzziele Gesundheit, Sicherheit und Grundrechte, einschließlich Demokratie, Rechtsstaatlichkeit und Umweltschutz, normativ gesetzt. Normative Mehrdeutigkeiten treten jedoch bei der Interpretation, Operationalisierung und Abwägung konkurrierender Werterealisierungen auf. Ebenso zeigen sie sich bei der Bestimmung von Restrisiken, die den Betroffenen aufgebürdet werden. Sie ergeben sich auch bei dem noch ungeklärten Verhältnis zwischen „AI safety" und dem Schutz von Grundrechten bei systemischen Risiken. Normative Mehrdeutigkeiten vermindern nicht nur die Effektivität der Risikoregulierung, sondern sie können auch selbst zur Ursache von Risiken werden.

Shaping European AI: the case of open source and digital sovereignty

Dr. Andreas Liesenfeld (Radboud University, The Netherlands)

Digital sovereignty, open source AI, constructive technology assessment, AI technology assessment

In light of the European Union's recent "Tech sovereignty package," open-source AI is increasingly positioned as a strategic solution to counter market monopolies and reduce technological dependencies. Within the European discourse, open source is frequently championed as a cornerstone for achieving digital sovereignty. So far, however, we have little evidence that open-source AI contributes to these goals. Drawing on four years of AI technology assessment as part of the European Open Source AI Index (www.osai-index.eu) project, this talk addresses contrasts between political narratives and empirical realities, as well as discusses the role of AI technology assessment as a (normative) force that actively shapes the future of generative AI in Europe (Rip et al. 1995).

This contribution highlights how AI technology assessment is essential for evaluating how much potential European AI has as a substantially different approach to, e.g. small, sustainable, local, no (!?) AI, or whether it is doomed to replicate a similar monolithic, centralized AI landscape, different in name only. Charting the potential for possible alternative, this form of AI technology

assessment facilitates a dialogue with the public and the providers of alternative technologies. Grounded in an anthropology of common concerns (Xiang 2022), this form of TA picks up lived worries and existential struggles that emerge in the public, and, in the form of a feedback loop, surveys whether emerging alternative approaches actually address these concerns, see community uptake, and make a difference in people's lives. We demonstrate this approach to technology assessment by examining the pursuit of digital sovereignty as a case study of thinking in alternative. Ultimately, we sketch a “TA of common concerns” that aims to enable community-driven forms of technology development and assessment.

KI als Methode der Forschung

Moderation: Dr. Caroline Dotter (OTH Regensburg, Deutschland)

Session 2d	Raum K004	09:00–11:00 Uhr

Wie diskutiert die Wissenschaftscommunity über genKI-Nutzung? Empirische Einblicke aus Bluesky-Daten zu Forschungspraktiken und Publikationskultur

Camille Landesvatter, Johanna Mehler, Angelina Sophie Dähms, Linda Nierling, Constanze Scherz, Leonie Seng, Marius Albiez, Ralf Schneider (ITAS, KIT, Deutschland)

Generative KI, Wissenschaftspraxis, Diskursanalyse, Social Media Analyse, Natural Language Processing

Die Nutzung von KI-basierten Anwendungen in der wissenschaftlichen Forschungs- und Arbeitspraxis ist inzwischen ebenso verbreitet wie diskutiert. Potenziale und Risiken für das Wissenschaftssystem werden längst nicht mehr nur innerhalb der Disziplinen – allen voran der Technikfolgenabschätzung (TA) – verhandelt, sondern auch außerhalb der Wissenschaft, in öffentlichen Diskursräumen.

Um nachzuvollziehen, wie der Einsatz von KI in der Wissenschaft aktuell diskutiert wird, untersuchen wir im Rahmen des Leibniz-WissenschaftsCampus „Digital Transformation of Research“: Welche Themen lassen sich im öffentlichen Diskurs zum Einsatz von genKI in der akademischen Forschung identifizieren? Welche Schritte in den jeweiligen Forschungsprozessen werden dabei besonders in den Blick genommen? Und welche Einschätzungen zu den Auswirkungen auf die wissenschaftliche Integrität lassen sich herausarbeiten?

Um einen direkten Einblick in den Diskurs zu erhalten, untersuchen wir die für die Forschungscommunity einschlägigen Räume auf Social Media, wie bspw. Bluesky. Dabei sammeln wir systematisch Postings, wobei die zugrundeliegende Suchstrategie Äußerungen aus drei Bereichen des Diskurses erfassen soll: wissenschaftliche Praxis (z. B. Ideengenerierung, Schreibprozess), Qualitätssicherung (z. B. Peer Review, Begutachtung, Science Governance), sowie den Publikationsprozess wissenschaftlicher Produkte (z. B. Qualität und Volumen von Veröffentlichungen). Anhand der gesammelten Posts und Threads analysieren wir den Diskurs mit Methoden der maschinellen Sprachverarbeitung und identifizieren dabei Akteure sowie Themen und Narrative in ihrer zeitlichen Entwicklung.

Die genannten Fragen bilden den Horizont des empirischen Projekts. Im Beitrag selbst fokussieren wir auf Suchstrategie, Exploration und Datenqualität des Materials sowie auf eine erste Einordnung deskriptiver Ergebnisse. Auf dieser Grundlage erörtern wir zukünftige Forschungsperspektiven, Anknüpfungspunkte und weiterführende Fragen für die TA.

KI in der Grundschule? Die dichte Beschreibung als methodischer Ausgangspunkt in der Technikfolgenabschätzung

Michael Decker[1], Barbara Bruno[2], Uta Hauck-Thum[3] ([1]TU München und Deutsches Museum, Deutschland, [2]Karlsruher Institut für Technologie, [3]Ludwig-Maximilians-Universität München)

Künstliche Intelligenz, Robotik, Dichte Beschreibung, Technikfolgenabschätzung, Mündliches Erzählen

Ziel einer Mensch-Technik-Interaktion ist es, dass sich Mensch und Technik optimal in einem Handlungszusammenhang ergänzen und damit das Handlungsziel besser erreichen. Die Beurteilung der Kooperation bedarf dann einer Beschreibung der Arbeitsteilung. Wer, Mensch oder künstlich intelligenter Roboter, macht was genau in dieser Kooperation?

Die fünf Stufen des autonomen Fahrens beschrieben z.B., was der Mensch und was das Fahrzeug beim „gemeinsamen Fahren" übernimmt und damit auch, ab welcher Stufe der Mensch nicht mehr seine volle Aufmerksamkeit auf das Fahren richten muss (Wienrich 2022). In Bezug auf die Pflege von älteren Menschen beschreibt Hülsken-Giesler (2025, 2020), dass gelingende Pflegearbeit als Interaktionsarbeit verschiedene Aspekte vereinen muss (i) Gefühlsarbeit (Umgang mit Gefühlen anderer), (ii) Emotionsarbeit (Umgang mit den Emotionen als Pflegende), (iii) Kooperationsarbeit (Kooperationsbeziehungen herstellen), (iv) Wissensbasiert (externe Evidenz) und körpernah, (v) Einzelfallbasiert (interne Evidenz, biographische Präferenzen und situativ). Er bewertet den Einsatz von Robotern in diesem Handlungszusammenhang als noch wenig passend („Mismatch") (2020, S. 148).

In diesem Beitrag wird auf der Basis einer dichten Beschreibung der Umsetzung des sprachdidaktischen Konzepts „Mündliches Erzählen" in der Grundschule beurteilt, welche Teilaufgaben ein kleiner, künstlich intelligenter, humanoider Roboter übernehmen kann. Ziel ist es, in der Kooperation von Lehrerin/Lehrer und Roboter ein insgesamt besseres Handlungsergebnis zu erzielen. In dem von der AIM Heilbronn geförderten Projekt arbeiten grundschulpädagogische Forschung und künstlich intelligente Robotikforschung interdisziplinär zusammen. Wobei diese dichte Beschreibung nicht nur für Orientierung bei der Entwicklung des Roboters und für Verständnis der Lehrerinnen und Lehrer sorgen kann, sondern auch für eine entwicklungsbegleitende Technikfolgenabschätzung hilfreich ist.

KI zwischen Kontrolle und Vertrauen: Technikphilosophische, epistemische und governancebezogene Perspektiven

Moderation: Constanze Scherz (Institut für Technikfolgenabschätzung und Systemanalyse, Deutschland)

Session 3a	Raum K001	13:45–15:15 Uhr

KI als nachmoderne Technik: Zur Technikfolgenabschätzung KI-basierter Technik

Prof. Dr. Jan Cornelius Schmidt (TH Darmstadt, Deutschland / European University of Technology (EUT+), Deutschland / Irland / Frankreich u.a.)

Prospektive TA, Methodologie der TA, nachmoderne Technik, Instabilität

Künstliche Intelligenz prägt Gegenwart und Zukunft spätmoderner Wissensgesellschaften wie kaum eine andere Technologie. Um gesellschaftliches Gestaltungswissen bereit zu stellen, ist die Technikfolgenabschätzung derzeit stark herausgefordert.

Im Folgenden sollen etablierte Zugänge der TA zur KI ergänzt werden, indem eine vertiefte Technikcharakterisierung vorgenommen wird, die den wissenschaftlich-technischen Kern von KI offenlegt und beurteilt. Der hier verfolgte Ansatz steht in der Tradition der Prospektiven Wissenschafts- und Technikfolgenabschätzung (ProTA, vgl. erstmals: Liebert/Schmidt 2010). Es wird gezeigt, dass wir Zeitzeugen der Entstehung eines gänzlich neuen Techniktyps sind: Nachmoderne Technik kann als epochale Ergänzung und Erweiterung der uns vertrauten klassisch-modernen Technik verstanden werden. Der Vortrag legt – auf Basis der Methoden von ProTA – den wissenschaftlich-technischen Kern nachmoderner Technik frei, wie er durch Künstliche Neuronale Netze induziert wird. Dieser Kern betrifft die technische Herstellung und trickreiche Nutzung von Instabilitäten – als Quelle von Selbstorganisation und dynamischer Komplexität. Damit bilden Instabilitäten die notwendige Bedingung für jene erwünschten Eigenschaften, die mit Großbegriffen verbunden werden, wie: Produktivität, Adaptivität/Flexibilität und Autonomie. Aus Perspektive der TA zeigt sich der Kern nachmoderner Technik als ambivalent: Die enormen Produktivitäts- und Performancegewinne werden erkauft durch Kontrollierbarkeits-, Verstehbarkeits- und Antizipations-Limitierungen.

Vor dem Hintergrund dieser ProTA-basierten Technikcharakterisierung werden schließlich Vorschläge präsentiert, die zeigen, wie mit dieser inhärenten Ambivalenz umgegangen werden könnte. Damit wird gleichzeitig deutlich, wie etablierte Zugänge der TA zu KI produktiv ergänzt und erweitert werden können.

Literatur

Liebert, W.; Schmidt, J.C. 2010: Towards a prospective technology assessment. In: Poiesis & Praxis 7. Jg., 99–116.

Epistemische Bedingungen von KI-Systemen und ihre praktischen Folgen

Stephan Altemeier (Philosophische Praxis NRW, Deutschland / Internationale Hochschule Erfurt, Deutschland)

Erkenntnistheorie, Praktische Philosophie, Technikfolgen, Künstliche Intelligenz, Episteme

Der aktuelle Entwicklungsstand sogenannter „Künstlicher Intelligenz" wirft die Frage auf, inwiefern hier tatsächlich „Intelligenz" vorliegt. Medial präsente Beispiele wie Chatbots oder autonomes Fahren sind letztlich hochkomplexer Output basierend auf immensem Daten-Input. Es handelt sich um Formen schwacher KI, deren „Wissen" aus mathematischen Theoremen und Wahrscheinlichkeitsintegration resultiert.

Da die Grenze dieser „Wissensproduktion" scheinbar nur in der Rechenleistung liegt, diskutiert die Forschung auch die Perspektive einer starken KI, die künftig ihre Umwelt selbstständig erfassen, ein Bewusstsein entwickeln und eigene Handlungsgründe generieren könnte.

Dieser Vortrag fokussiert daher die erkenntnistheoretischen Implikationen schwacher KI und deren Auswirkungen auf die Alltagsanwendung. Die Dringlichkeit dieses Anliegens unterstreicht die aktuelle Meldung aus dem Jahr 2026 über Anthropics KI-Modell „Mythos": Es gilt potenziell als hochgefährlich, da es rasend schnell Sicherheitslücken in digitalen Infrastrukturen aufdeckt und die zugehörigen Exploits direkt mitliefert.

Hier drängt sich Horkheimers Erkenntnis auf, dass die „Anwendung der Theorie auf das Material nicht nur innerszientistischer, sondern zugleich ein gesellschaftlicher Vorgang" ist. Die Entwicklung von „Mythos" demonstriert eindrücklich, dass die epistemischen Bedingungen schwacher KI zentral dafür sind, wie wir das Verhältnis von Gesellschaft und KI-Systemen denken müssen – ein Verständnis, das zwingend für einen reflektierten KI-Einsatz erforderlich ist.

Synthetische Daten, kalibriertes Vertrauen: Verantwortung und Kontrolle in der KI-gestützten Qualitätskontrolle

Alena Bischoff (Universität Augsburg, Deutschland)

Synthetische Anomaliedaten, kalibriertes Vertrauen, Trust-by-Design, Understandability, Bias

KI-gestützte Qualitätskontrolle steht vor einem Datenparadox: Seltene, sicherheitsrelevante Fehler sollen zuverlässig erkannt werden, treten aber zu selten auf, um KI-Systeme mit realen Anomaliedaten zu trainieren. Die Erhebung realer Trainingsdaten ist zudem technisch aufwendig und rechtlich voraussetzungsreich. Synthetische Anomaliedaten versprechen hier eine Abhilfe, erzeugen jedoch neue Unsicherheiten in der Vertrauensfrage: Sie mindern die Datenknappheit, bergen aber auch Risiken falscher Evidenz, eines Verlusts des Realitätsbezugs, von Bias, der Umgehung von Einwilligungserfordernissen (consent circumvention) und von Governance-Lücken (Whitney & Norman, 2024; Lee, 2025; McCormack et al., 2025; Shafiee & Farid, 2026).

Der Beitrag entwickelt am interdisziplinären Projekt „SynthEthik" einen Ethics-by-Design-Ansatz, der ein „kalibriertes" Vertrauen in der Datenpipeline verankert, das weder in ein Over- noch Under-Trust umschlägt. Leitende These ist, dass Vertrauen nicht erst an der Nutzungsoberfläche entsteht, sondern entlang der Pipeline: von der Datenerhebung über synthetische Anomaliegenerierung, Training und Validierung bis hin zur Visualisierung, Konfiguration durch DomänenexpertInnen und Entscheidungen im Fehlerfall einer konkreten Anwendung. Gefragt wird, welche

Informationen, Kontrollmöglichkeiten und Verantwortungsstrukturen QualitätsprüferInnen benötigen, um KI-Empfehlungen auf synthetischer Datenbasis prüfen, übernehmen oder zurückweisen zu können.

Theoretisch verbindet der Beitrag die Human-Factors-Forschung zu Use, Misuse und Disuse (Parasuraman & Riley, 1997), das Drei-Ebenen-Modell dispositionalen, situativen und erlernten Vertrauens (Hoff & Bashir, 2015), nutzerzentrierte Transparenz bzw. Understandability (Werz et al., 2025) sowie das Delegationsmodell des Deutschen Ethikrats (2023). Ziel ist ein TA-Kriterienraster für Trust-by-Design: Provenienz, Plausibilität, Unsicherheit, Validierung, Override, Eskalation und Verantwortungszuweisung.

KI in therapeutischen und gesundheitsnahen Berufen – Schwerpunkt: interaktive KI

Moderation: Constanze Scherz (Institut für Technikfolgenabschätzung und Systemanalyse, Deutschland)

Session 3b	Raum K002	13:45–15:15 Uhr

Die perfekte Simulationsperson? Zur Automatisierung von Simulation und Feedback in der medizinischen Aus- und Weiterbildung

Elena Loevskaya

Künstliche Intelligenz, Simulationspersonen, Feedback, Medical Education

Simulationspersonenprogramme sind ein etablierter Bestandteil der medizinischen Aus- und Weiterbildung. Simulationspersonen werden in Kommunikations- und Interaktionstrainings, praktischen Prüfungen sowie interprofessionellen Lehrsettings eingesetzt. Dabei übernehmen sie nicht nur die Darstellung standardisierter Rollen, sondern häufig auch eine zweite Funktion: die Rückmeldung subjektiver Gesprächs- und Beziehungserfahrungen im Rahmen strukturierter Feedbackprozesse.

Mit der Entwicklung generativer KI entstehen zunehmend KI-basierte Simulationspersonen sowie automatisierte Trainings- und Feedbacksysteme. Aktuelle Arbeiten diskutieren den Einsatz großer Sprachmodelle als Simulationspersonen und untersuchen deren Potenzial hinsichtlich Skalierbarkeit, Standardisierung und Verfügbarkeit.

Der Beitrag greift diese Entwicklungen auf, verschiebt jedoch den Fokus der Debatte. Statt zu fragen, ob Simulationspersonen durch KI ersetzt werden können, wird untersucht, welche Funktionen Simulationspersonen im Ausbildungskontext überhaupt erfüllen und inwiefern diese automatisierbar sind. Dabei wird zwischen der Simulationsfunktion – der Darstellung einer Rolle – und der Feedbackfunktion unterschieden, die auf subjektiver Wahrnehmung, Beziehungserfahrung und Reflexion basiert.

Im Zentrum steht die Frage, welche Vorstellungen von Kommunikation, Lernen und Bewertung sichtbar werden, wenn Rückmeldung und Reflexion zunehmend als automatisierbare Prozesse verstanden werden. Ziel des Beitrags ist eine technikfolgenorientierte Analyse der Frage, welche Aspekte simulationsbasierter Aus- und Weiterbildung durch KI ergänzt, transformiert oder ersetzt werden können und welche Formen zwischenmenschlicher Erfahrung dabei neu bewertet werden.

Ethische und soziale Aspekte von sozialer Robotik – der Roboter Navel aus Perspektive jüngerer und älterer Menschen mit sensorischen, kognitiven und emotionalen Einschränkungen

Prof. Dr. Sonja Haug, Dr. Caroline Dotter, Felicitas Braun, Vanessa Mücke, Sarah Nachreiner, Dr. Debora Frommeld, Prof. Dr. Karsten Weber (OTH Regensburg, Deutschland)

Soziale Robotik, Mensch-Roboter-Kommunikation, vulnerable Nutzergruppen, uncanny valley, Technikakzeptanz

Robotik wird zunehmend für den Einsatz in Einrichtungen im Pflegebereich diskutiert. Mit dem Einsatz von Robotern gehen jedoch verschiedene ethische und kommunikative Herausforderungen einher, insbesondere, wenn Roboter mit (älteren) Menschen mit Sinneseinschränkungen und/oder kognitiven Einschränkungen interagieren. Das Projekt **Soziale Robotik im Praxistest (SoRoP) – Eine explorative Studie mit (älteren) Menschen mit Sinnes- und Kommunikationseinschränkungen** erforscht mit drei Untersuchungsgruppen (ältere Menschen im stationären Setting, jüngere Menschen mit Autismus-Spektrum-Störung im teilstationären Setting, Vergleichsgruppe allgemeine Bevölkerung) die Mensch-Roboter-Interaktion anhand des sozialen Roboters „Navel" (navel robotics München). Untersucht werden Unterschiede in den Bewertungen des Roboters in Bezug auf Mimik, Gestik und Sprache abhängig von der Untersuchungsgruppe sowie Technikakzeptanz aufseiten der Befragten. Weiterhin wird untersucht, inwieweit der Roboter als unheimlich empfunden wird (Stichwort: „uncanny valley"). Die Feldstudie basiert auf einem Mixed-Methods-Forschungsdesign, bestehend aus einer explorativen Feldbeobachtung von Interaktionen, einem Gruppengespräch und einer Dual-Mode-Umfrage. Besonderes Augenmerk wird auf (forschungs-)ethische Aspekte gelegt.

Es werden erste quantitative Ergebnisse aus der Befragung präsentiert, ergänzt mit qualitativen Ergebnissen aus der Beobachtungsstudie. Mit allen Untersuchungsgruppen fanden erfolgreiche Kommunikationen statt. Die Kommunikationsfähigkeiten des Roboters wurden als gut eingeschätzt und der Roboter nicht als unheimlich empfunden. Analysiert werden Unterschiede in den Bewertungen der wahrgenommenen Lustigkeit, Natürlichkeit sowie verbalen und nonverbalen Kommunikation. Diskutiert werden ethische und soziale Aspekte und es werden Schlussfolgerungen für den zukünftigen Einsatz im Sozial- und Gesundheitsbereich abgeleitet.

Social Delusion? – Die Ambivalenz des Einsatzes von GenAI für die mentale Gesundheit

Maximilian Zhang (Universität Tübingen, Deutschland)

KI für Mentale Gesundheit, Cybertherapie, KI-Therapie, Mensch-KI-Kommunikation

Weltweit ist jede siebte Person von einer psychischen Störung betroffen (WHO, 2025). Doch die meisten von ihnen haben keinen Zugang zu professioneller Behandlung (Ebd.). Seit generative KI allgegenwärtig ist, wenden sich immer mehr Menschen, insbesondere Jugendliche, mit mentalen Problemen wie Stress, Kummer oder Depressionen an KI (Handelsblatt, 2026; Tagesschau, 2026). Einerseits berichten viele Nutzer:innen von positiven Erfahrungen, andererseits bereitet der therapeutische Einsatz von KI den Expert:innen große Sorgen. Denn unter der oberflächlichen Funktionalität lauern große Probleme wie emotionale Abhängigkeit und die Isolation von der Realität. Dies ist noch problematischer, wenn es um Voice Interfaces geht, da die sehr menschenähnliche künstliche Stimme eine größere soziale Präsenz erzeugt und menschliche Dialoge imitiert. In dem eingereichten Vortrag soll zunächst aus interdisziplinärer, theoretischer Perspektive erörtert wer-

den, wie der Kommunikationsprozess zwischen Mensch und voice-basierter KI aussieht und funktioniert und welche Rolle die KI dabei einnimmt. Anschließend werden die Ergebnisse der ersten Analyse der empirischen Daten einer qualitativen Studie vorgestellt. In dieser wurden die Nutzer:innen- und Expert:innenperspektiven zum Einsatz generativer KI zur Unterstützung der psychischen Gesundheit abgefragt. Am Ende des Vortrags wird das Thema noch einmal systematisch reflektiert und die Chancen und Risiken sowie die Implementierung und ethische Verantwortlichkeiten diskutiert und zusammengefasst.

KI und die Zukunft der Demokratie

Moderation: Maximilian Schultz (OTH Regensburg, Deutschland)

Session 3b	Raum K003	13:45–15:15 Uhr

KI-Nutzung durch Politiker*innen – wo liegen die Grenzen?

Prof. Dr. Alma Kolleck (TH Würzburg Schweinfurt, Deutschland)

Generative Künstliche Intelligenz, KI, Politik, Reden, Authentizität

Generative Künstliche Intelligenz transformiert die politische Kommunikation grundlegend, indem sie die Erstellung von Inhalten automatisiert und damit traditionelle Vorstellungen von Autorschaft infrage stellt [1;2]. Aktuell manifestiert sich diese Transformation exemplarisch im Fall von Mario Voigt, dem vorgeworfen wird, Reden, Zeitungsbeiträge und Social-Media-Posts KI-generiert zu haben. Die daraus resultierende mediale Debatte dient als Brennglas für eine Aushandlung über die Grenzen von KI im politischen Raum. So schreibt beispielsweise die taz,[1] dass eine KI-generierte Pressemitteilung toleriert werden könne, es hingegen „erbärmlich" sei, eine Rede zum Holocaust-Gedenktag ohne „eigene Gedanken" mit KI zu erstellen. Die ZEIT urteilt: „Wem Worte egal sind, der sollte kein Politiker werden."[2] Der Stern titelt: „Fragwürdiger Umgang mit KI-Texten: Ist Mario Voigt ein Blender?".[3]

Der Fall Voigt verdeutlicht, dass die Akzeptanz von KI-Nutzung stark von der Art und Intention von Texten sowie der zugeschriebenen „Authentizität" abhängt [1]. Es rücken Fragen ins Zentrum wie: Wo verorten verschiedene Akteure (in Journalismus, Politik, Öffentlichkeit) die Grenzen zulässiger KI-Nutzung? Wie wird diese Grenzziehung begründet und welche Ansprüche an Autorschaft werden dabei zugrunde gelegt? Der Vortrag zeigt auf, dass der Fall Voigt nicht nur ein lokales Polit-Ereignis ist, sondern exemplarisch für die notwendige Neudefinition von politischer Autorschaft, „Authentizität" und Verantwortung im Zeitalter algorithmischer Textproduktion steht [3]. Ziel ist es, die Kriterien freizulegen, anhand derer im politischen Raum „menschliche" Autorschaft als unabdingbar angesehen wird und wann „synthetische" Kommunikation akzeptabel erscheint. Zur Beantwortung dieser Fragen wird eine qualitative Inhaltsanalyse der medialen Berichterstattung und Kommentierung zum Fall Mario Voigt durchgeführt. Anhand von einem konkreten Fall werden somit grundlegend neu zu verhandelnde „Grenzen des Generierbaren" herausgearbeitet.

Literatur

[1] Beacken/Wooley (2026): Generative AI, Emerging Technology, and Political Communication: Examples, Methods of Study, and Implications for Democracy. In: Merlo/Liebowitz (Eds.): Ethical AI and Data Science: Building Trustworthy and Transparent Systems. ERC, pp. 171 – 184, DOI: 10.1201/9781003657804-11.

[2] Labbé et al. (2025): ChatGPT as speechwriter for the French presidents. Digital Scholarship in the Humanities, 40 (2), pp. 549 – 561, DOI: 10.1093/llc/fqaf032.

[3] Lucke/Sander (2025): A few Thoughts on the Use of ChatGPT, GPT 3.5, GPT-4 and LLMs in Parliaments: Reflecting on the results of experimenting with LLMs in the parliamentarian context. Digit. Gov.: Res. Pract. 6, 2, Article 27, 21 pages. DOI: 10.1145/3665333.

Zeitungsquellen

1. https://taz.de/KI-Reden-von-Mario-Voigt/!6186148/
2. https://www.zeit.de/politik/deutschland/2026-06/kuenstliche-intelligenz-politik-mario-voigt-rede
3. https://www.stern.de/politik/deutschland/mario-voigts-fragwuerdiger-umgang-mit-ki-texten--ist-er-einblender--37527232.html

Digitale Tools, KI und Demokratie – Ergebnisse einer STOA-Studie

Davy van Doren, Bert Droste-Franke, Ulrich Hofmann (IQIB, Deutschland)

Künstliche Intelligenz, digitale Partizipationsinstrumente, politische Entscheidungsprozesse, UN-Nachhaltigkeitsziele, Bürgerbeteiligung

Die zunehmende Integration Künstlicher Intelligenz (KI) in digitale Partizipationswerkzeuge verändert die Schnittstelle zwischen Bürger*innen und politischen Entscheidungsprozessen grundlegend (Macintosh, 2004). Im Kontext der UN-Nachhaltigkeitsziele bietet KI große Potenziale, aber birgt ohne reflexive Gestaltung jedoch das Risiko einer gesellschaftlichen und institutionellen Orientierungslosigkeit (UN DESA, 2022).

Bereits wurde argumentiert, dass KI-Tools kein rein technologisches Surrogat für tiefgreifende Beteiligungsprozesse sein sollen, sondern dass ein demokratischer Mehrwert erst dann entsteht, wenn eine strikte funktionale Kohärenz zwischen dem Partizipationsziel, dem Prozess, den Software-Spezifikationen und einer wirksamen menschlichen Aufsicht gewahrt bleibt (Floridi et al., 2018).

Dieser Beitrag verknüpft die Ergebnisse einer STOA-Studie zu E-Demokratie mit den Anforderungen nachhaltiger Governance. Eine Untersuchung von 94 Tools zeigt unter anderem, dass KI-Funktionen die Bürgerbeteiligung im Politikzyklus quantitativ und qualitativ ausweiten könnten. Gleichzeitig drohen durch stochastische Fehler, algorithmische Verzerrungen und mangelnde digitale Barrierefreiheit neue Exklusionsmechanismen, die den integrativen Anspruch von SDGs untergraben. Der Beitrag wird einige Orientierungshilfen liefern für die TA, wie der Wandel von digitaler Technologie resilient und rechenschaftspflichtig gestaltet werden kann.

Literatur

Floridi, L., Cowls, J., Beltrametti, M., et al. (2018). AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. Minds and Machines, 28(4), 689–707.

Macintosh, A. (2004). Characterizing e-participation in policy-making. System Sciences, 2004. Proceedings of the 37th Annual Hawaii International Conference on, 5–11.

United Nations Department of Economic and Social Affairs (UN DESA). (2022). E-Government Survey 2022: Digital Transformation in the Future of Public Administration. United Nations.

DIGIT(AI)ZED EMPOWERMENT

Evelyn Hriberšek (www.eurydike.org, Deutschland)

Digitale Gewalt, GenAI, Empowerment, Science-Fiction, Kunst

Wie können wir Zukünfte verantwortungsvoll mitgestalten angesichts patriarchaler Machtstrukturen von Tech-Bros, digitaler Gewalt gegen FLINTA* und dem Einfluss von GenAI? Die XR-Pionierin Evelyn Hriberšek gibt Einblick in ihr künstlerisches Schaffen und spannt einen Bogen von Mythologie über Science Fiction zur Technikfolgenabschätzung, um exemplarisch Handlungsoptionen aufzuzeigen.

2014 löste die Medienkritikerin Anita Sarkeesian mit einer Analyse von weiblichen Stereotypen in Games einen weltweiten Shitstorm mit Vergewaltigungs- und Morddrohungen aus: #Gamergate wurde zu einer männerdominierten Kampagne gegen Frauen in der Videospielindustrie und wurde 2017 durch die #MeToo-Debatte überholt.

2020 thematisierte Terre de Femmes mit der Kampagne #UnhateWoman (sprachliche) Gewalt gegen Frauen, die millionenfach "gehört, geliked und gefeiert" wird und sich im Digitalen als Hatespeech fortsetzt. Als eine Feministin den Rapper Fler mit dem Hashtag markierte, drohte er mit realer Gewalt, bei einer anderen Kritikerin setzte er ein Kopfgeld auf Instagram aus, um sie zum Schweigen zu bringen.

Frauenfeindlichkeit, Silencing und Gewaltverherrlichung erfahren seit 2022 durch GenAI und die Möglichkeiten von Deepfake eine neue Dimension. Plattformbetreibende entziehen sich zugunsten von „Kreativität“ und „Redefreiheit“ jeglicher Verantwortung: Digitale Gewalt dient als Geschäftsmodell. In diesem Kontext spielen AI-Glasses – häufig genutzt von sog. Pick-Up-Artists – eine akute Rolle: Wearables, mit denen Kinder und/oder Frauen heimlich gefilmt und im Netz zur Schau gestellt werden.

Noch gibt es keine Gesetze zum Schutz gegen digitale Gewalt, es existieren jedoch kreative Gegenpositionen. So empowert die Mixed Reality Experience EURYDIKE das Publikum und macht gleichzeitig Silencing erlebbar. Interaktiv bringen User*innen eine verstummte, real erfahrbare Sci-Fi-Welt zum Klingen: Stellvertretend für alle Ungehörten und Ungesehenen wird EURYDIKE so eine Stimme verliehen.

Normative Orientierung der TA in Zeiten von KI

Moderation: Dr. Linda Nierling; Dr. Philipp Frey (Institut für Technikfolgenabschätzung und Systemanalyse, Deutschland)

Session 3d	Raum K004	13:45–15:15 Uhr

Zwischen Versprechen und Verantwortung: KI, Nachhaltigkeit und die normative Neuverortung der Wissenschaft im digitalen Wandel

Silke Niehoff[1], Grischa Beier[1,2] ([1]RIFS Forschungsinstitut für Nachhaltigkeit | am GFZ, Deutschland, [2]Universität Potsdam)

Digitalisierung, Künstliche Intelligenz, Nachhaltigkeit, Twin Transformation, Rolle der Wissenschaft

Der gegenwärtige technologische Wandel, insbesondere die rasche Verbreitung Künstlicher Intelligenz (KI), stellt die wissenschaftliche Praxis vor grundlegende Fragen. Digitale Transformation trägt nicht automatisch zu nachhaltiger Entwicklung bei. KI wird oft als technologisches Allheilmittel für Nachhaltigkeitsprobleme positioniert, ohne dass Rebound-Effekte und systemische Risiken hinreichend mitgedacht werden. Der Betrieb von KI-Systemen ist mit erheblichen Ressourcenaufwänden verbunden und steht damit in Konkurrenz zu gesellschaftlichen Dekarbonisierungszielen (Freitag et al. 2021). Große Sprachmodelle, die zunehmend wissenschaftliche und private Routinen prägen, können das Verständnis dessen, was als „nachhaltig" gilt, verändern und wichtige Aushandlungsprozesse über Gerechtigkeit und ökologische Grenzen beschneiden (Bush et al. 2025).

Der Beitrag nimmt diese Konstellation zum Ausgangspunkt, um die Rolle der Wissenschaft im Kontext der Twin Transformation kritisch zu beleuchten. Er fokussiert auf die Frage, unter welchen normativen Voraussetzungen wissenschaftliche Akteure KI als Transformationsinstrument einsetzen, kritisch begleiten oder begrenzen sollten.

Literatur

Bush, A.; Aksoy, M.; Pauly, M.; Ontrup, G. (2025): Choosing a Model, Shaping a Future: Comparing LLM Perspectives on Sustainability and its Relationship with AI. In: Conference on Empirical Methods in Natural Language Processing, p. 17330–17341. DOI: 10.18653/v1/2025.findings-emnlp.939.

Freitag, C.; Berners-Lee, M.; Widdicks, K.; Knowles, B.; Blair, G. S.; Friday, A. (2021): The real climate and transformative impact of ICT: A critique of estimates, trends, and regulations. In: Patterns 2 (9), S. 100340. DOI: 10.1016/j.patter.2021.100340.

Moralisch ambitionierte TA in Zeiten der Krise? – Ein Diskussionsbeitrag

Dennis Mandwurf (VDI/VDE-IT, Deutschland)

Krise, TA, KI, Moral Ambition, Optimismus

„Wir erleben eine Zeitenwende. Das bedeutet: Die Welt danach ist nicht mehr dieselbe wie die Welt davor." Das Zitat ist interessant, weil es nichts darüber aussagt, wie „die Welt danach" sein wird – besser, schlechter? Klar ist nur, die Welt wird anders sein. Eine Vision hatte Kanzler Scholz nicht;

dagegen vertrat seine Vorgängerin mit dem Satz „Wir schaffen das!“ noch einen naiven Optimismus. Und so bestimmt eine dauerhafte Krisenstimmung Entscheidungsverfahren und Meinungsbildung. Hinzu kommt, dass der Einsatz Künstlicher Intelligenz wenig Visionäres anzubieten hat – KI ist nicht optimistisch oder ambitioniert.

Und so stellt sich die Frage, welche Rolle TA als Beratungsinstrument in der „Zeitenwende“ einnimmt: Kann TA nicht nur Vision Assessment betreiben, sondern selbst Vision Provider sein? Oder anders gefragt: Sollte TA den gesellschaftlichen Wandel konnotieren und durch die Betonung intendierter Folgen einen (Fortschritts-)Optimismus vertreten (ohne selbstverständlich auf Nebenfolgen hinzuweisen)? Falls ja, welche Auswirkungen hätte eine Positionierung der TA auf Organisation und Gemeinschaft, institutionelle Einbindung, Transfer und Kommunikation, Methoden, Finanzierungsmodelle sowie Nachwuchsgewinnung?

Anhand des Appells zur Moral Ambition des niederländischen Historikers und (Technik-)Optimisten Rutger Bregman soll diskutiert werden, welchen Beitrag TA in Zeiten eines gesellschaftlichen Wandels in Krisenstimmung hin zu einer besseren Welt leisten kann.

Literatur

Bogner, Alexander 2021: Politisierung, Demokratisierung, Pragmatisierung, Paradigmen der Technikfolgenabschätzung im Wandel, in: Böschen et al. (Hrsg.): Technikfolgenabschätzung, Baden-Baden: Nomos: 43-58.

Dusseldorp, Marc 2014: Technikfolgenabschätzung zwischen Neutralität und Bewertung, in: APuZ, 6-7: 25-30.

Grunwald, Armin 2008: Technik und Politikberatung – Philosophische Perspektiven, Frankfurt a.M.: Suhrkamp.

Grunwald, Armin 2025: Technikfolgenabschätzung – Einführung, 4. Aufl., Baden-Baden: Nomos.

Anschlussfähigkeit statt Rollen: Technikfolgenabschätzung im Kontext von KI-getriebener Transformation

Janine Gondolf, Paulina Dobroc (Karlsruher Institut für Technologie, Deutschland)

Rollen der TA, TA-Forschung, Methodenentwicklung, Praxisrelevanz, Orientierungskompetenz

Der rasche Fortschritt der künstlichen Intelligenz (KI) verstärkt den Bedarf an evidenzbasiertem Orientierungswissen, dessen Bereitstellung zu den Kernaufgaben der Technikfolgenabschätzung (TA) gehört. Gleichzeitig gerät die TA selbst unter Druck, da im Zuge der Transformation Wissensgrundlagen und Vorgehensweisen infrage gestellt werden. Der Beitrag nimmt diese Spannung zum Ausgangspunkt, um die Orientierungskompetenz der TA neu zu denken.

Die TA beschreibt ihre Tätigkeiten traditionell als Rollen (z. B. Beraterin, Analytikerin, Moderatorin) und macht damit ihr Ringen um eine wissenschaftlich fundierte und normativ wirksame Position sichtbar. Aufbauend auf einer Literaturanalyse zeigt der Beitrag die Vielzahl der Rollenbeschreibungen und der darauf aufbauenden Erklärsysteme. In qualitativ angelegten Workshops mit TA-Praktiker:innen wurden diese Rollenbilder diskutiert und anschließend konkrete Projektsituationen rekonstruiert und vergleichend analysiert. Die Ergebnisse zeigen, dass die Orientierungsleistung der TA situativ durch Passungen zwischen Problemkonstellationen, Akteurserwartungen und Interaktionsformen entsteht. Rollenbeschreibungen fungieren dabei primär als nachträgliche Deutungen.

Vor diesem Hintergrund wird der Begriff der „Anschlussfähigkeit“ als analytische und reflexive Perspektive eingeführt. TA-Arbeit wird dabei als situatives, kollaboratives Herstellen von Relevanz und Orientierung im Kontext verstanden. Dadurch wird die Frage nach der Expertise der TA im Kontext KI-getriebener Transformationen geschärft. Diese lässt sich entlang zweier Dimensionen präzisieren: der Auswahl und Strukturierung von Problemstellungen („Was erforscht die TA?“) sowie der Gestaltung von Interaktions- und Wissensformaten („Wie forscht die TA?“).

Der Beitrag zeigt, wie TA Orientierung herstellt und leistet damit einen Beitrag zur konzeptionellen Neujustierung der TA im Kontext KI-getriebener Transformationsprozesse.

Deepfakes zwischen Vertrauen, Verantwortung und Kontrollverlust

Moderation: Dr.-Ing. Jutta Jahnel (Institut für Technikfolgenabschätzung und Systemanalyse, Deutschland)

Jutta Jahnel[1], Dana Mahr[1], Maria Pawelec[2], Anna Louban[2], Murat Karaboga[3], Christian Geminn[4], Leon Francis[4]

Session 4a	Raum K001	15:30–17:00 Uhr

Täuschend echt? Deepfakes zwischen Vertrauen, Verantwortung und Kontrollverlust

Ziel der Session

Die Session möchte erfahrbar machen, wie sogenannte Deepfakes – also durch künstliche Intelligenz erzeugte oder manipulierte audiovisuelle Inhalte – sowohl gesellschaftliche Orientierungsprozesse herausfordern als auch kriminelles Missbrauchspotenzial bergen. Im Mittelpunkt steht nicht nur die technische Funktionsweise, sondern die Frage, wie Betroffene, Institutionen und gesellschaftliche Akteure mit dieser neuen Form von Unsicherheit umgehen. Ziel ist es, die gesellschaftlichen, ethischen und institutionellen Folgen zu beleuchten und den Beitrag der Technikfolgenabschätzung zu diesen Debatten sichtbar zu machen.

Kurzbeschreibung

Deepfakes sind längst kein Randphänomen digitaler Kultur mehr. Sie beeinflussen politische Diskurse, journalistische Praktiken und individuelle Wahrnehmungen von Realität. Besonders gravierend zeigen sich ihre Folgen im Bereich sexualisierender Deepfakes, die häufig Frauen und marginalisierte Gruppen betreffen und neue Herausforderungen für Datenschutz, Persönlichkeitsrechte, demokratische Prozesse und Strafverfolgung schaffen. Zugleich eröffnen Deepfakes neue Ausdrucksformen – etwa in Kunst, Satire oder Erinnerungskultur.

Die Session bringt eine Betroffene sowie Akteurinnen und Akteure aus Wissenschaft, Bildung, Medienpraxis, Technikfolgenabschätzung und Jugendsozialarbeit zusammen, um zu diskutieren, wie Vertrauen und Verantwortung in einer durch KI geprägten Medienumgebung neu ausgehandelt werden. Sie fragt, welche normativen und rechtlichen Antworten entstehen und umzusetzen sind, wie unterschiedliche gesellschaftliche Gruppen reagieren und welche neuen Formen der Verantwortung erforderlich sind.

Ablauf und methodisches Format Die Session ist dialogisch und interaktiv angelegt:

1. Einführung
   Kurze Einführung in Thematik und Ablauf durch die Moderation.
2. Demonstration der technischen Möglichkeiten zur Erstellung von Deepfakes von Lasse Kohlmeyer, Hasso-Plattner-Institut (HPI) und Prof. Matthias Wölfel, Hochschule Karlsruhe (HKA)
3. Moderierte Panel-Diskussion mit Lasse Kohlmeyer (HPI), Prof. Matthias Wölfel (HKA), Felicia Ewert (Autorin, Referentin), Patrik Stemmer (Jugendsozialarbeiter) und Steffen Albrecht (ITAS, TAB).
4. Öffnung der Panel-Diskussion für das Publikum
5. Zusammenführung der Ergebnisse

[1]Karlsruher Institut für Technologie, Institut für Technikfolgenabschätzung und Systemanalyse (ITAS)

[2]Eberhard Karls Universität Tübingen, Internationales Zentrum für Ethik in den Wissenschaften (IZEW)

[3]Fraunhofer-Institut für System- und Innovationsforschung (ISI)

[4]Universität Kassel, Fachgebiet Öffentliches Recht, IT-Recht und Umweltrecht

KI in therapeutischen und gesundheitsnahen Berufen – Schwerpunkt: administrative KI

Moderation: Dr. Diana Schneider, Dr. Nils Heyen (Fraunhofer-Institut für System- und Innovationsforschung, Deutschland)

Session 4b	Raum K002	15:30–17:00 Uhr

AI-Assisted Clinical Documentation: Efficiency Gains and Physician Experience - A Scoping Review

Sophie Püchel, Agnes Kandlbinder, Ralf Jox (Centre Hospitalier Universitaire Vaudois (CHUV), Institut des Humanités en Médecine, Université de Lausanne, Schweiz)

Clinical documentation, hospital discharge letters, time efficiency, LLMs

Clinical documentation represents a substantial and growing share of physicians' workload in hospital care and has been linked to reduced patient interaction, increased time pressure, and a higher risk of burnout. Physicians now spend nearly half of their working time on administrative tasks and electronic health records (Sinsky et al. 2016). Over time, discharge letters have evolved from simple communication tools between physicians into multifunctional documents serving legal, administrative, reimbursement, educational, and patient-information purposes, substantially increasing documentation demands. Large Language Models are increasingly promoted to reduce documentation burden and improve efficiency, but real-world impact remains unclear.

Emerging evidence suggests that AI-assisted documentation can reduce the time required for clinical documentation. However, these gains may not necessarily decrease physicians' overall administrative workload and may instead shift it to additional tasks. This reflects the concept of the "time paradox," whereby efficiency gains do not automatically translate into a perceived reduction in workload (Rosa 2005).

This scoping review highlights a potential mismatch between AI-related time savings in clinical documentation and physicians' perceived administrative burden. We examine both objective outcomes (e.g. efficiency gains) and subjective experiences (e.g. perceived workload), and studies linking the two.

Quantitative time savings may not lead to increased patient interaction or improved physician well-being, but may instead contribute to temporal compression. Meaningful improvements require broader organizational changes beyond technological optimization alone.

Literature

Rosa H. (2013). Social Acceleration: A New Theory of Modernity. New York: Columbia University Press.

Sinsky C. (2016). Allocation of Physician Time in Ambulatory Practice: A Time and Motion Study in 4 Specialties. Ann Intern Med, 165(11):753-760.

Transformation der Arbeitspraktiken durch Generative AI: Analyse der Nutzungsmöglichkeiten für die Psychotherapie in der Schweiz

Roman Heggli (Berner Fachhochschule, Schweiz)

Generative AI; Psychotherapie; Affordance-Theorie; digitale Transformation

Die rasante Entwicklung von Generative AI (GenAI) eröffnet neue Möglichkeiten zur Entlastung selbstständiger Psychotherapeut:innen in der Schweiz. Diese Masterarbeit untersucht, wie GenAI in nicht-therapeutischen Aufgaben wie Dokumentation, Kommunikation und Qualitätssicherung unterstützen und dadurch Ressourcen für die psychotherapeutische Arbeit freisetzen kann.

Die empirische Grundlage bilden qualitative, semi-strukturierte Interviews mit elf Psychotherapeut:innen. Die Analyse erfolgt theoriegeleitet anhand der Affordance-Theorie (Hutchby, 2001). Diese Perspektive zeigt, dass das Potenzial von GenAI nicht allein in der Technologie liegt, sondern erst in der konkreten Nutzung entsteht.

Die Ergebnisse zeigen Potenzial vor allem bei Berichten für externe Stellen, der Transkription und Zusammenfassung von Erstgesprächen sowie der Bearbeitung von Therapieplatzanfragen. In diesen Bereichen kann GenAI entlasten und Ressourcen für die Therapie freisetzen. Für komplexere Aufgaben wie diagnostische Einschätzungen oder Supervision wird GenAI hingegen nur vereinzelt als sinnvoll wahrgenommen. Skepsis, mangelndes Vertrauen, fehlende Anwendungskompetenzen und Datenschutzbedenken begrenzen die Nutzung.

Der Nutzen von GenAI entsteht durch intuitive, individualisierbare und datenschutzkonforme Tools, die auf den Arbeitsalltag selbstständiger Psychotherapeut:innen in der Schweiz zugeschnitten sind. Entscheidend sind zudem Unterstützung, Praxisbeispiele und kollegialer Austausch, damit Vertrauen und Anwendungskompetenzen aufgebaut werden können. Für Psychotherapeut:innen liegt das Potenzial in der gezielten Entlastung administrativer Prozesse, während Berufsverbände und Entwickler:innen die Integration durch verständliche Informationsangebote, nutzungsnahe Anwendungen und begleitende Implementierungsangebote unterstützen können.

Literatur

Hutchby, I. (2001). Technologies, Texts and Affordances. Sociology, 35(2), 441–456. https://doi.org/10.1177/S0038038501000219

Zwischen Vertrauen und Wandel: Gesellschaftliche Perspektiven auf KI

Moderation: Franziska Hauer (OTH Regensburg, Deutschland)

Session 4c	Raum K003	15:30–17:00 Uhr

Jenseits der theoretischen Neugierde: Die wachsende Relevanz philosophischer Reflexion für gesellschaftliche Verständigungs- und Gestaltungsprozesse

Martina Philippi (Helmut-Schmidt-Universität / Universität der Bundeswehr Hamburg, Deutschland)

Deutungsautonomie, Ambiguität, Resilienz, Ideologie, Dark Enlightenment

Lange galt philosophische Reflexion aus Alltagsperspektive als verzichtbar, akademisch und weltfremd. Klassische philosophische Fragen wie *Was ist der Mensch?* oder *Was heißt Denken?* schienen die Lebenswirklichkeit weder in gutem noch in schlechtem Sinne zu verändern.

Dies hat sich mit der Einbettung Künstlicher Intelligenz in Alltags- und professionelle Routinen geändert. Diese Fragen stehen nun im Zentrum gesellschaftlicher Verständigung über die Gestaltung und Regulierung einer neuartigen Technologie mit globalem gesellschaftlichem, politischem, ökologischem und sogar existenziellem Impact: In welcher Welt wollen wir leben? Was unterscheidet uns von den Maschinen?

Doch der gesellschaftlichen Verständigung droht abhanden zu kommen, was seit jeher die Grundlage philosophischer Reflexion war: Wir brauchen ein begriffliches Instrumentarium, Problembewusstsein und Ambiguitätstoleranz, um uns diese(n) Fragen zu stellen. Das auszubilden und aussichtsreich zu verfolgen, fordert Zeit und Gestaltungsspielraum. Diese Ressourcen werden zum einen bedroht durch Narrative und Ideologien, die sich im Kontext des ‚Dark Enlightenment' wiederfinden: Akzelerationismus, Solutionismus, behavioristische Verhaltenssteuerung. Zugleich ist es ausgerechnet das Denken selbst, das mit der fortschreitenden Etablierung von KI in Teilschritte fragmentiert und automatisiert oder in standardisierte Prozesse überführt wird.

Mein Beitrag thematisiert die Bedeutung von Reflexionsräumen für Deutungsautonomie und Resilienz gegenüber Ideologien: Nur indem wir uns im gesellschaftlichen Diskurs über Menschenbilder und den Wert unseres Denkens verständigen, schützen wir uns vor Narrativen und Ideologien, die nicht nur den Menschen als kommerzielle oder politische Ressource instrumentalisieren und Reflexions- und Verständigungsprozesse korrumpieren, sondern dies in die Technologien und somit die Praxen einbetten, die unseren Alltag, unser gesellschaftliches Miteinander und letztlich politische Macht prägen.

Zwischen Hoffnung und Sorge: Einstellungen zum technologischen Wandel im 25-Länder-Vergleich

Jakob Kaiser, Andreas Neus, Carolin Kaiser (Nuremberg Institute for Market Decisions, Deutschland)

Technologischer Wandel, soziale Ungleichheit, Arbeitsmarkt, Mehrländervergleich

Neue Technologien wie Künstliche Intelligenz werden häufig als Treiber tiefgreifender gesellschaftlicher Veränderungen wahrgenommen, verbunden mit Hoffnungen auf Wohlstand ebenso wie mit Sorgen vor sozialen und ökonomischen Verwerfungen. Um technologische Wandlungsprozesse sozialverträglich zu gestalten und empirisch fundierte Orientierung zu bieten, ist ein systematisches Verständnis der Einstellungen in der Bevölkerung wichtig.

Wir präsentieren Ergebnisse einer repräsentativen quantitativen Befragung in 25 Ländern (je ca. 1.000 Befragte, N ≈ 25.000) zu den erwarteten wirtschaftlichen und sozialen Folgen des aktuellen technologischen Wandels. Die Bewertung des technologischen Wandels variiert deutlich mit dem Wohlstandsniveau der Herkunftsländer: Befragte aus wohlhabenden Ländern (z. B. Deutschland, Schweiz) blicken überwiegend pessimistisch auf neue Technologien, während Befragte aus Schwellenländern (z. B. Indien, Südafrika) signifikant optimistischer sind. Unabhängig vom wirtschaftlichen Status des Landes erwartet jedoch in fast allen Ländern eine Mehrheit, dass neue Technologien die Arbeitslosigkeit erhöhen werden. Rund 40 % der Befragten weltweit gehen zudem davon aus, dass der derzeitige technologische Wandel soziale Ungleichheit verschärft, d.h. Vorteile für wohlhabendere und Nachteile für ärmere Bevölkerungsgruppen schafft.

Die Ergebnisse verdeutlichen eine global verbreitete Ambivalenz im Umgang mit neuen Technologien: Selbst dort, wo Technologien als Chance für wirtschaftlichen Aufschwung gelten, dominieren Sorgen vor Arbeitsplatzverlusten und wachsender Ungleichheit. Für die Technikfolgenabschätzung liefern diese Befunde eine empirische Grundlage, um öffentliche Erwartungen und Befürchtungen differenziert in Politikberatung und Technikgestaltung einzubeziehen.

Alltagshelfer oder Apokalypse? | Was junge Menschen auf dem Lande über KI denken

Marius Albiez, Steffen Bauer, Markus Winkelmann, Constanze Scherz, Julian Sann (Karlsruher Institut für Technologie, Deutschland)

Nachhaltige Entwicklung, Zukunftsbilder, Futures Literacy, TA

Ende 2022 veröffentlichte OpenAI den Chatbot ChatGPT. Die niedrigschwellige Bedienung und Nutzung durch textbasierte Prompts führte innerhalb weniger Wochen zur millionenfachen Nutzung und löste eine öffentliche Debatte über Chancen und Risiken von KI für die Gesellschaft aus. Der Boom des maschinellen Lernens und Deep Learnings war nun auch in der breiten Öffentlichkeit angekommen. Damit stiegen auch die Erwartungen an die Technikfolgenabschätzung (TA), fundierte Einschätzungen zu einer sich rapide entwickelnden Technologie zu geben. Ein Arbeitsfeld ist dabei die Reflexion von KI im Kontext des Leitbilds Nachhaltiger Entwicklung und die Untersuchung, wie KI die Chancen auf ein gutes Leben heute und in Zukunft beeinflusst.

Kinder und Jugendliche spielen eine zentrale Rolle, da sie von entsprechenden zukünftigen Entwicklungen vergleichsweise lange betroffen sind und sie eine Brücke zwischen heutigen und zukünftigen Generationen bilden. Gut begründete TA-Analysen sollten dabei vielfältige Methoden, Datenquellen und Perspektiven einbeziehen. Im Kontrast zu den in der TA häufig behandelten

städtischen Räumen (z. B. Smart-City), sehen die Autor*innen analytischen Nachholbedarf bei der Berücksichtigung ländlicher und nicht-urbaner Räume.

Vor diesem Hintergrund soll im Vortrag folgender Frage nachgegangen werden: „Wie blicken junge Menschen auf dem Lande auf KI-bezogene Entwicklungen?“ Dabei werden wünschenswerte und negativ konnotierte Zukunftsvorstellungen analysiert – von Hausaufgaben-Hilfen oder Pflege-Robotern bis zu befürchteten Sicherheitsrisiken oder der Veränderung menschlicher Beziehungen. Als empirische Grundlage für den Vortrag dienen Erkenntnisse aus dem Projekt „Mobiles Futurium“ (2023–2026). Der Beitrag fußt auf über 100 leitfadengestützten Interviews mit Schüler*innen an sechs Schulen in unterschiedlichen Bundesländern in Deutschland (Mai 2024 – Jan. 2026). Neben der Vorstellung unserer Analysen werden auch Empfehlungen für politisch Entscheidende formuliert.

Funktionale Realität des Seins (FRS): Ein Rahmen zur Analyse von KI-Risiken jenseits der Bewusstseinsfrage

Moderation: Dr. Michael Boronowsky; Dr. Bert Droste-Franke (Institut für qualifizierende Innovationsforschung & -beratung, Deutschland)

Session 4d	Raum K004	15:30–17:00 Uhr

Impuls zu FRS

Dr. Michael Boronowsky, Dr. Bert Droste-Franke

Two Wrongs Don't Make a Right: Über- und Unter-Anthropomorphisierung von LLMs

Reinhard Heil, Wolfgang Eppler (Karlsruher Institut für Technologie, Deutschland)

Generative KI, Anthropomorphisierung, Bias

Während viel über die Gefahren der Anthropomorphisierung großer Sprachmodelle (LLMs) geschrieben wurde, wird nicht ausreichend beachtet, dass es ebenso riskant sein kann zu wenig zu anthropomorphisieren. In diesem Beitrag argumentieren wir, dass sowohl eine übermäßige als auch eine unzureichende Anthropomorphisierung („Computerisierung") erhebliche, wenn auch unterschiedliche epistemische und praktische Bedrohungen darstellen und sich diese Bedrohungen zudem wechselseitig verstärken.

Einerseits können wir, indem wir LLMs zu viel Menschlichkeit zuschreiben, zu der Annahme verleitet werden, dass diese Systeme ähnlich denken wie Menschen oder gar empfindungsfähig oder moralisch verantwortlich sind. Andererseits behandeln wir LLMs so, als wären sie herkömmliche Software-Tools. Wir „computerisieren" also LLMs, indem wir davon ausgehen, dass die von LLMs produzierten Aussagen das gleiche epistemische Gewicht haben wie Ergebnisse von regelbasierten Softwaresystemen und akzeptieren sie ohne ausreichende Prüfung. Anders gesagt, wir gehen mit der Anthropomorphisierung nicht weit genug, da wir uns mit LLMs nicht so auseinandersetzen, wie wir es im Idealfall mit menschlichen Gesprächspartnern tun würden, deren Aussagen wir hinterfragen, untersuchen und interpretieren.

Ironischerweise riskieren wir, indem wir LLMs via „Computerisierung" bestimmte anthropomorphe Eigenschaften (wie Fehlbarkeit, Voreingenommenheit oder Kontextabhängigkeit) absprechen, ihre Genauigkeit zu überschätzen und ihre Fähigkeit, Fehler zu machen, zu unterschätzen. In unserem Beitrag zeigen wir, dass sich weder die Anthropomorphisierung als Korrektiv der Computerisierung eignet noch die Computerisierung als Korrektiv der Anthropomorphisierung. Vielmehr besteht die Gefahr, dass sich deren negative Effekte wechselseitig verstärken.

Maschinenrational und Schließung der Welt. Technikfolgenabschätzung als Projekt der Aufklärung

Marcel Krüger[1], Benjamin Rathgeber[2] ([1]Karlsruher Institut für Technologie, Deutschland, [2]Hochschule für Philosophie München, Deutschland)

Anthropomorphismus, Technomorphismus, Zweck-Mittel-Verkehrung, Determinismus, Freiheit

Dys- wie utopischen Entwürfen vom Einsatz derjenigen Technologien, die unterschiedslos mit Künstlicher Intelligenz betitelt werden, mangelt es an einem wesentlichen Reflexionsmoment: Im jeweiligen angeblichen Potential von KIs unterstellen sie ihnen Eigenschaften, die sich bestimmten Selbstbeschreibungen des Menschen und Vorstellungen von deren bruchlosen Übersetzung in Kalküle verdanken, aber nicht mehr als Projektionen sind. Problematisch ist denn auch nicht der Einsatz von KIs überhaupt, sondern a) ein Missverständnis ihrer Funktionsweise und Fähigkeiten und damit verbunden die Verfehlung sinnvoller Einsatzzwecke sowie b) die Tendenz, in Hinsicht auf die Selbst- und Weltverhältnisse der Menschen einer verabsolutierten Grammatik der Digitalisierung selbst zum Opfer zu fallen. Die Konsequenzen lassen sich bereits heute beobachten, wenn sich in den Wissenschaften und in Politik die Pluralität der Weltbeschreibungen einem Primat der Digitalisierbarkeit unterordnet und die Freiheit des Menschen Algorithmen weichen soll: Wie sich der Mensch versteht, welche Vermögen ihn und was ein gutes (Zusammen-)Leben auszeichnen, wird zunehmend von der Logik der Maschine her gedacht, die prinzipielle Offenheit menschlicher Selbstauslegung durch ein Technologieparadigma abgelöst. Die Welt der Menschen schließt sich.

Im Vortrag werden wir diesen doppelten Shift vom Anthropomorphismus der Technologie zum Technomorphismus der Welterschließung und von der Verkehrung von Zweck und Mittel nachzeichnen und auf dieser Basis einen Vorschlag entfalten, wie TA in der Tradition der Aufklärung diese Tendenz kritisch begleiten und Orientierung für eine offene Ordnung geben kann.

Literatur

Feige, D.M. (2025): Kritik der Digitalisierung. Technik, Rationalität und Kunst

Mühlhoff, R. (2025): Künstliche Intelligenz und der neue Faschismus

Nosthoff, A.-V. (2026): Kybernetik und Kritik. Eine Theorie digitaler Regierungskunst

Trawny, P. (2015): Technik. Kapital. Medium. Das Universale und die Freiheit

MITTWOCH
23. September 2026

KI fährt voraus – Autonome Mobilität gestalten: Eine Zukunftswerkstatt aus Sicht der Technikfolgenabschätzung (Workshop)

Moderation: Ulrike Scorna; Vanessa Mücke (OTH Regensburg, Deutschland)

Session 5a	Raum K003	09:00–10:30 Uhr

Wann sind wir da? (Wirklich!) Selbstfahrende Fahrzeuge nach 15 Jahren.

Torsten Fleischer (Karlsruher Institut für Technologie, Deutschland)

Autonomes Fahren, Regulierung, Innovationsgestaltung, Mobilität, KI

Auch wenn es schon in den Jahrzehnten zuvor zahlreiche Aktivitäten zur Umsetzung selbstfahrender Fahrzeuge gegeben hat: 2011 war ein wichtiges Jahr in deren Entwicklung. Google erwarb zwei Robotik- und KI-Startups und formte damit den Kern dessen, was heute als Waymo bekannt ist. Im gleichen Jahr veröffentlichte einer der damals leitenden Entwickler, Sebastian Thrun, in der New York Times einen Meinungsbeitrag, der als expectation statement für autonomes Fahren (AF) bis in die Gegenwart fortwirkt. Heute fahren Robotaxis in mehreren Städten in den USA im Regelbetrieb, chinesische Unternehmen sind in ihrem Heimatland aktiv und expandieren in den Nahen Osten und zunehmend auch nach Europa. TA und benachbarte Fächer haben diese Entwicklungen seit vielen Jahren begleitet. Dabei blieben die Einschätzungen immer ambivalent zwischen Sorge um Verkehrszuwachs und Machtverschiebungen in der Daseinsvorsorge einerseits und Hoffnung auf neue, nachhaltigere Mobilitätsdienstleistungen und erweiterte Teilhabeoptionen andererseits.

Der Beitrag versucht sich an einer Bestandsaufnahme. Er will der Frage nachspüren, warum es in Europa bisher keine solchen Fahrzeuge im Regelbetrieb gibt und dabei vor allem einen Blick auf institutionelle Bedingungen werfen. Eine solche Perspektive bietet, neben Einsichten für die Mobilitätspolitik, auch Hinweise für das Thema der Tagung. AF in der heutigen Form ist nur durch die Nutzung von KI möglich – und diese Entwicklung ist angesichts der großen allozierten Ressourcen in dem Feld weiterhin hoch dynamisch, denn hier stehen immer noch unterschiedliche System-„Philosophien“ und Modelle im Wettbewerb. Zugleich sind KI-Implementierungen in Fahrzeugen unmittelbar sicherheitsrelevant, in ihren Wirkungen öffentlich sichtbar und aufgrund der Regelungsdichte im Straßenverkehr eine quasi prototypische regulatorische Herausforderung. Am Beispiel des AF lassen sich die Möglichkeiten und Herausforderungen gesellschaftlicher Gestaltung von KI „live“ studieren.

Workshop

Zwischen Vertrauen und Wandel: Gesellschaftliche Perspektiven auf KI

Moderation: Franziska Hauer (OTH Regensburg, Deutschland)

Session 5b	Raum K002	09:00–10:00 Uhr

Biografische Brüche als sensible Phasen für den Einfluss von Künstlicher Intelligenz auf zwischenmenschliche Beziehungen

Anna Pauls, Fabian Fischer, Ulrike Bechtold (Österreichische Akademie der Wissenschaften, Österreich)

KI-Gefährten, zwischenmenschliche Beziehungen, Lebensphasen

Verschiedene KI-Anwendungen erfüllen zunehmend Bedürfnisse nach sozialem Austausch oder werden dafür angepriesen: Von personalisierten Chatbots (etwa mit „Persönlichkeit“ konfiguriert) über KI-Avatare, KI-Agenten bis hin zu dezidierten KI-Gefährten (AI Companions). In unserem Beitrag werden wir eine Systematisierung der Bedeutung verschiedener Lebenssituationen und biografischer Brüche für die Nutzung von KI-Beziehungs-Gegenüber und deren möglichen Auswirkungen auf offline-Beziehungen vornehmen. Wir diskutieren und kategorisieren Lebensphasen, die mit Brüchen in Biografien verbunden sind. Dazu gehören Reifeprozesse (z. B. Adoleszenz), aber auch private, berufliche, gesellschaftliche und gesundheitliche Kontexte. Wir untersuchen, inwieweit sensible Phasen, die mit solchen Brüchen einhergehen, eine besondere Beeinflussbarkeit bedeuten und das Beziehungsgeflecht der Betroffenen verändern – dies könnte sowohl unter Nutzer:innen als auch zwischen Nutzer:innen und Nicht-Nutzer:innen von KI der Fall sein. Wir werden anhand bereits existierender Forschung darlegen, welche Bedeutung dies für die Nutzung von KI-Gefährten und ähnlichen Anwendungen hat und wie eine Typisierung sowie eine Systematisierung anhand verschiedener Lebensphasen und Arten zwischenmenschlicher Beziehungen aussehen kann. Beispiele für Brüche, die eine besondere Beeinflussbarkeit des Beziehungsgeflechts bedeuten könnten, sind im Arbeitskontext das Einarbeiten in einen neuen Beruf, neuen Arbeitsplatz, das Ausscheiden vom Arbeitsplatz/Berufsleben; im privaten Kontext die Geburt des ersten Kindes, der Verlust eines nahestehenden Menschen, der Auszug der Kinder, Pensionseintritt; oder im sozialen Kontext Diskriminierung anlässlich des Alters, geänderter Gesundheitssituation oder Genderstatus.

KI, betrügerische Deepfakes und die Industrialisierung von Vertrauensmissbrauch

Anna Louban, Maria Pawelec (Internationales Zentrum für Ethik in den Wissenschaften, Deutschland)

Betrug, Deepfakes, Vertrauensmissbrauch, generative KI, Kriminalität

KI ermöglicht eine Industrialisierung von Vertrauensmissbrauch durch automatisierte, personalisierte und skalierbare Täuschungspraktiken. KI-generierte oder -veränderte Bild-, Ton- und Videoinhalte manipulieren Betroffene systematisch. Vulnerabel sind nicht nur Menschen mit bestimmten Einschränkungen, psychologischen Charakteristika oder in Situationen erhöhter Unsicherheit. Durch fortgeschrittene KI-Technologien und ausgefeilte Betrugsmaschen können immer mehr Personen zu Betroffenen werden.

Der Beitrag untersucht KI- und v.a. Deepfake-gestützte Betrugspraktiken, die sich an Privatpersonen richten, als Ausdruck eines strukturellen Mechanismus der Skalierung von Vertrauensmissbrauch. Deepfake-gestützte Betrugsfälle haben massiv zugenommen und treten in immer mehr Bereichen auf. Anhand ausgewählter Deepfake-Betrugsarten (z.B. health fraud, Reise-Scam) analysiert der Beitrag, wer von solchen Betrugsformen betroffen ist, wie sie funktionieren und wie sie in bestehende gesellschaftliche Zusammenhänge eingebettet sind.

Der Beitrag argumentiert, dass KI-gestützter Vertrauensmissbrauch weder rein individualisiert noch isoliert zu betrachten, sondern sozio-technisch verankert ist. Schließlich basiert er auf technischen Fortschritten im Bereich generativer KI und zunehmend auf der Verschränkung digitaler Daten-, Plattform- und Kommunikationsinfrastrukturen. Dazu zählen v.a. datengetriebene Werbeökonomien, Data-Broker-Systeme und Callcenter-Strukturen, die gemeinsam die Grundlage für skalierbare personalisierte Täuschung bilden. ‚Betrugserfolge' basieren daher nicht primär auf individuellen Fehlentscheidungen, sondern insbesondere auf technisch und ökonomisch verschränkten Macht- und Infrastrukturen.

Unsere Analyse dieser strukturellen Zusammenhänge ermöglicht eine Einordnung der Risiken und ihrer gesellschaftlichen Tragweite. Damit trägt der Beitrag zur Orientierung bei und schafft Grundlagen für politische, gesellschaftliche und individuelle Handlungsoptionen.

Interdisziplinäre Perspektiven auf KI

Moderation: Prof. Dr. Karsten Weber (OTH Regensburg, Deutschland)

Session 5c	Raum K001	09:00–11:00 Uhr

Was sind die Beiträge von generativer KI zu extremistischer Kommunikation? Erkenntnisse der Technikfolgenabschätzung zur zivilen Sicherheit

Christian Büscher, Alexandros Gazos, Tim Röller (KIT, Deutschland)

Radikalisierungsdynamiken, soziotechnische Systeme, Automatisierte Sinnproduktion, Kommunikationseskalation, Risikoprävention

Künstliche Intelligenz entwickelt sich rasant. Insbesondere kommerzielle Anwendungen können überzeugende Texte, Bilder oder Videos erstellen. Dadurch sind sie potenziell in der Lage, zu sozialer Kommunikation beizutragen. Dies hat gesellschaftliche Folgen, die auch das Thema „Zivile Sicherheit“ und „Extremismusprävention“ betreffen.

In aktuellen soziologischen Diskursen wird die Beteiligung von Maschinen an der sozialen Kommunikation diskutiert, als Frage einer Qualifikation, die bislang nur Menschen vorbehalten war (Baecker), als prädiktiver Generator interessanter, informativer, zuverlässiger oder unterhaltsamer Inhalte (Esposito) oder als Akteur in einem Imitationsspiel (Dickel). Es stellt sich deshalb die Frage nach der Funktion der Beiträge von generativer KI in der Kommunikation. Vor allem die Erzeugung extremistischer Inhalte (rassistisch, sexistisch, Gewalt und Hass fördernd), die in unendlich vielen Varianten in den alltäglichen Kommunikationsfluss eingespeist und zum Teil für „bare Münze“ genommen werden, fordert eine umfassende Gefahrenabschätzung heraus.

In diesem Beitrag wollen wir anhand der Kommunikationstheorie, die eine dreifache Selektion von Information, Mitteilung und Verstehen annimmt, eine Heuristik zur Besprechung von aktuellen Szenarien der extremistischen Kommunikation entwickeln, die wir in der wissenschaftlichen Aufarbeitung dieser Thematik vorfinden.

Dahingehend sollte die weitere Entwicklung von KI von der Technikfolgenabschätzung begleitet und fortwährend beobachtet werden, um politische Entscheidungsträger:innen zu orientieren und der Prävention gezielte Interventionen zu ermöglichen. Für die präventiven Möglichkeiten generativer KI gilt es hingegen, die damit verbundenen praktischen, ethischen und rechtlichen Herausforderungen herauszuarbeiten.

KI und digitale Transformation am Arbeitsplatz – Praxisbericht eines Szenarioprozesses

Martin J. Kirstgen[1], Titus Udrea[2], Ewa Dönitz[1], Simone Kimpeler[1] ([1]Fraunhofer Institut für System- und Innovationsforschung, Deutschland, [2]AK Wien – Büro für digitale Agenden, Österreich)

KI, Arbeitswelt, Szenarien, Foresight, Horizon Scanning

Der breite Einsatz von KI und der für Laien nicht übersehbare Fortschritt führen zu Unsicherheit und Orientierungslosigkeit. Insbesondere auch im Kontext Arbeit. Arbeitnehmende aller Qualifikationsstufen befürchten, dass KI menschliche Tätigkeiten verdrängt. Neben dem Blick auf mögliche technische Optimierungen bedarf es daher eines Ansatzes, der die Herausforderungen und

Chancen der Arbeitswelt der Zukunft mit KI und digitaler Transformation aus der Perspektive von Arbeitnehmenden und Mitbestimmungsakteuren beleuchtet.

In der hier vorgestellten Studie wurde dies mit einem partizipativen Szenarioprozess untersucht. Der Prozess umfasste ein Horizon Scanning zur Identifizierung von relevanten Einflussfaktoren auf die Zukunft der Arbeit im Zeitalter von KI, sowie die Szenario-Entwicklung und Analyse von zukünftigen Herausforderungen. Dabei wurden mit Expert:innen und Stakeholdern zunächst vier Szenarien für die Arbeitswelt im Jahr 2040 entwickelt. Daraus wurden sowohl szenariospezifisch als auch übergreifend Chancen und Risiken für Arbeitnehmende in unterschiedlichen Tätigkeitsbereichen diskutiert und Handlungsbedarfe abgeleitet. Auf dieser Basis konnten strategische Handlungsoptionen für die Interessenvertretung der Arbeitnehmenden ausgearbeitet werden. Diese nähern sich dem Thema KI und Arbeit aus menschenzentrierter Perspektive und leisten einen Beitrag zur Orientierung für Arbeitnehmende, Mitbestimmungsakteure, politische Entscheider:innen und Unternehmen in einem volatilen Umfeld.

Im Beitrag wird die Relevanz der Fragestellung für eine Institution der Arbeitnehmerinteressen dargelegt und die für den Prozess konzipierte Methodik vom Horizon Scanning zum Szenario-Sprint vorgestellt. Dabei werden Herausforderungen beleuchtet, die sich aus den verschiedenen Schutzinteressen der vielfältigen Stakeholder im Themengebiet Arbeit und KI ergeben haben und ausgewählte Handlungsoptionen zur Gestaltung der KI-basierten Arbeitswelt im Interesse von Arbeitnehmenden vorgestellt.

Verantwortung, Vertrauen, Orientierung? Eine interviewbasierte TA-Perspektive auf KI im Verteilnetzbetrieb

Clemens Alexander Ackerl (Karlsruher Institut für Technologie, Deutschland)

Künstliche Intelligenz, betriebliche Praktiken, Vertrauen, Verantwortung, kritische Infrastrukturen

Fragen der Verantwortung und des Vertrauens stehen im Zentrum aktueller Debatten um KI in kritischen Infrastrukturen. Der Beitrag untersucht diese Fragen am Beispiel des KI-gestützten Verteilnetzbetriebs und veränderter betrieblicher Praktiken. Verteilnetze rücken im Zuge der Energiewende ins Zentrum einer sozio-technischen Transformation: Dezentrale Einspeisung, Bidirektionalität, Elektromobilität, Wärmepumpen und Speicher sind Herausforderungen in Mittel- und Niederspannungsnetzen. KI-Verfahren versprechen bessere Prognosen, Engpasserkennung und Entscheidungsunterstützung – kurz: eine Antwort auf steigende Komplexität.

Auf Basis von sechs Interviews mit Verteilnetzbetreibern in Deutschland analysiert der Beitrag, wie diese KI wahrnehmen und welche Erwartungen und Unsicherheiten entstehen. Die Analyse zeigt: KI ist kein technisches Upgrade, sondern eine epistemische Rekonfiguration von Arbeit im Netzbetrieb. Neben Erfahrungswissen, regelbasierten Verfahren und lokaler Netzkenntnis treten probabilistische Vorhersagen und algorithmische Optimierung.

Dadurch entstehen Reibungen zwischen Netzexpertise und algorithmischen Wissensformen. Im Verteilnetz werden Entscheidungen häufig unter Bedingungen unvollständiger Daten, lokaler Besonderheiten und eingeschränkter Netzzustandstransparenz getroffen. Vertrauen wird zu einem zentralen Mechanismus, durch den Netzbetreiber Unsicherheiten algorithmischer Entscheidungsunterstützung bearbeiten. Es wird als normative und relationale Erwartung verstanden, die in sozio-technischen Interaktionen unter Unsicherheit entsteht, Komplexität reduziert und Handeln koordiniert.

Der Beitrag liefert einen empirisch fundierten Impuls zur Integration von KI-Systemen in betriebliche Praktiken kritischer Infrastrukturen. Für die TA zeigt sich, dass Orientierung im Umgang mit KI nicht allein durch technische Leistung entsteht, sondern durch die Verknüpfung mit Expertise, organisationalen Zuständigkeiten und geklärten Verantwortlichkeiten.

Ethische Führung in der KI-gestützten Produktion

Dr. Sebastian Rosengrün (Universität Augsburg, Deutschland)

KI in der Produktion, Industrie, Führungsverantwortung, Mitbestimmung

Künstliche Intelligenz verändert industrielle Produktion nicht nur technisch, sondern auch sozial. Sie greift in Aufgaben, Abläufe und Verantwortlichkeiten ein – und verändert damit den Arbeitsalltag derjenigen, die mit ihr arbeiten. Der Vortrag stellt Ergebnisse einer qualitativen Studie vor, die am KI-Produktionsnetzwerk der Universität Augsburg mit Industriepartnern durchgeführt wurde. Im Zentrum steht die Frage, wie Beschäftigte und strategisch Verantwortliche diese Veränderungen wahrnehmen: Wo erleben sie Entlastung, Effizienzgewinne und bessere Wissensverfügbarkeit – und wo entstehen Unsicherheit, Kontrollbedenken, Kompetenzverluste oder neue Formen der Verantwortungsdiffusion?

Die leitende These: KI ist in der Produktion kein bloßes Werkzeug, das bestehende Prozesse beschleunigt. Sie rekonfiguriert Arbeit. Aufgaben werden neu verteilt, Schnittstellen verschieben sich, Erfahrungswissen wird anders sichtbar, Entscheidungen werden datenförmiger vorbereitet, Verantwortlichkeiten müssen neu geklärt werden. Genau darin liegt die ethische Führungsaufgabe. Führung im KI-Zeitalter heißt nicht nur, Systeme einzuführen und Akzeptanz zu erzeugen – sie heißt, die sozialen Folgen technologischer Veränderung aktiv zu gestalten.

Daraus folgt ein erweitertes Verständnis ethischer Führung: KI-Einführung ist als organisationaler und normativer Veränderungsprozess zu begreifen. Dazu gehören transparente Kommunikation, frühzeitige Beteiligung, klare Verantwortungsstrukturen, realistische Kompetenzentwicklung und ein sensibler Umgang mit Ängsten, Kontrollfragen und Machtverschiebungen. Ethische Führung zeigt sich dort, wo Beschäftigte nicht bloß „mitgenommen", sondern als Mitgestaltende ernst genommen werden.

Verantwortlicher KI-Einsatz in der Produktion ist damit weniger eine Frage abstrakter Prinzipien als eine konkrete Führungsaufgabe: gute Arbeit, Vertrauen und Verantwortlichkeit unter veränderten technischen Bedingungen möglich zu machen.

Funktionale Realität des Seins (FRS): Ein Rahmen zur Analyse von KI-Risiken jenseits der Bewusstseinsfrage

Moderation: Dr. Michael Boronowsky; Dr. Bert Droste-Franke (Institut für qualifizierende Innovationsforschung & -beratung, Deutschland)

Session 5d	Raum K002	10:00-11:00 Uhr

Feldwahrnehmung als leibliche Qualität in KI-bezogenen TA-Prozessen: Über das nicht-automatisierbare Komplement zur Künstlichen Intelligenz

Roland, J. Schuster (FH des BFI Wien, Österreich)

Feldwahrnehmung, Propriozeption, Rangdynamik, Omega-Position, Interventionsforschung

Partizipative Technikfolgenabschätzung (TA) konfrontiert Forschende mit komplexen Stakeholder-Dynamiken, die sich nicht allein kognitiv erfassen lassen. Der vorliegende Beitrag argumentiert, dass die leibliche Wahrnehmungskapazität der Forschenden – hier als Feldwahrnehmung bezeichnet – ein zentrales, nicht-automatisierbares Komplement zu KI-gestützten Analysemethoden darstellt.

Aufbauend auf der Klagenfurter Interventionsforschung (Schuster 2024, Kap. 4) wird das rangdynamische Modell von Raoul Schindler (2016) herangezogen, um die Omega-Position als strukturellen Ort der Feldwahrnehmung in Gruppen zu identifizieren: Die Person in dieser Position registriert abgespaltene Gruppenaffekte nicht analytisch, sondern leiblich – als somatische Resonanz, die dem reflektierten Verstehen vorausgeht (vgl. Bion 1962; Gendlin 1981). Diese Qualität ist zugleich das zentrale Instrument und die zentrale Vulnerabilität der Forschenden.

In KI-bezogenen TA-Prozessen (vgl. Grunwald 2010) gewinnt diese Dimension besondere Relevanz: Während KI die datenanalytischen Kapazitäten erweitert, bleibt die leibliche Feldwahrnehmung – das Spüren von Machtasymmetrien, verdeckten Konflikten und emotionalen Untergründen in Stakeholder-Gruppen – eine genuin menschliche Qualität, die algorithmisch weder repliziert noch ersetzt werden kann.

Der Beitrag führt den kinästhetischen Selbstrahmen als propriozeptive Technologie ein: eine leiblich verankerte Fixierung, die es Forschenden ermöglicht, ihre Feldwahrnehmung aufrechtzuerhalten, ohne in den wahrgenommenen Dynamiken aufzugehen. Diese innere Fixierung fundiert die äußere formale Fixierung (Schuster 2024, S. 104) leiblich.

Literatur

Bion, W. R. (1962). Learning from Experience. Heinemann.

Gendlin, E. T. (1981). Focusing. Bantam.

Grunwald, A. (2010). Technikfolgenabschätzung – eine Einführung (2. Aufl.). Edition Sigma.

Schindler, R. (2016). Das lebendige Gefüge der Gruppe. Psychosozial-Verlag.

Schuster, R. J. (2024). Interventionswissenschaft. transcript.

Des Kaisers neue Kleider

Jaro Krieger-Lamina (FH Technikum Wien, Österreich)

Gesellschaft, Demokratie, Menschenrechte, Künstliche Intelligenz

Künstliche Intelligenz (KI) verändert vieles. KI soll als saubere Technologie erscheinen, obwohl Kühlung und Strom einen enormen Fußabdruck hinterlassen. Sie soll Psychotherapie ersetzen – allerdings nur für jene, die sich keine Therapeut:innen leisten können. KI wird für Rechtsberatung, Effizienz in der Verwaltung, neue Forschungsansätze und als Lösung für den Klimawandel angepriesen. Manche erwarten von ihr Antworten auf die Frage nach dem Sinn des Lebens oder nutzen sie, um mit Verstorbenen zu „sprechen". Andere befürchten, KI werde eines Tages intelligenter als der Mensch. Aber können KI-Systeme das überhaupt?

Nein. Dennoch findet der Wandel statt. Die Narrative und Versprechen der KI-Hersteller haben reale Folgen: Sie lenken Investitionen, Verwaltung, sogar Verteidigung, setzen zunehmend auf KI, und Tech-Konzerne gewinnen an Macht. Staaten lagern Aufgaben an private Tech-Unternehmen aus und machen sich von diesen abhängig. Dahinter zeigen sich ideologische Tendenzen hin zu Technokratie und zu totalitären und faschistischen Idealen.

Die „guten" und „bösen" KI-Systeme verraten letztlich mehr über menschliche Wünsche und Ängste als über die Technik selbst. Gleichzeitig geraten die tatsächlichen gesellschaftlichen Veränderungen oft aus dem Blick.

Konsens in der TA? Während manche TA-Praktiker:innen KI für unverzichtbar halten, sehen andere sie als brandgefährlich. Dabei sollte TA die Leistungsfähigkeit einer Technologie realistisch bewerten können, um ihre Chancen und Risiken abwägen und Handlungsoptionen für eine wünschenswertere Zukunft aufzeigen zu können. Die normative Grundlage der TA mag angesichts von Weltrettungsfantasien banal erscheinen, ist aber dennoch wichtig: Es geht auch bei KI um Menschenrechte, um die Art des Zusammenlebens – und möglicherweise um die Zukunft der Menschheit. Voraussetzung dafür ist, den Versprechungen der Technologiekonzerne mit nüchternem Blick zu begegnen und des Kaisers neue Kleider als solche zu erkennen.

KI- und Robotik-Anwendungen bewerten: Nachhaltigkeits- und Ethikkriterien im Dialog

Moderation: Dr. Simon Hirsbrunner; Dr. Lou Brandner; Aline Franzke; Prof. Dr. Thomas Potthas, (Internationales Zentrum für Ethik in den Wissenschaften, Deutschland)

Session 5e	Raum K004	09:00–10:30 Uhr

KI und Klimawandel – ein gutes Team? Gedanken zur Rolle der Wissenschaft in dieser Beziehung

Ulrike Bechtold (Österreichische Akademie der Wissenschaften, Institut für Technikfolgen-Abschätzung, Österreich)

Klimawandel, Nachhaltigkeit, KI, Partizipation, Verantwortung

Mit Blick auf die Wechselbeziehungen zwischen Nachhaltigkeit und Technologie werde ich zentrale Narrative im Zusammenhang mit „globaler Erwärmung und KI" identifizieren und analysieren. Darauf aufbauend skizziere ich Aspekte von gesamtgesellschaftlichen Entscheidungen und welche Effekte verkürzte Narrative haben, die einen einfachen Ausweg aus den aktuellen Krisen Klimawandel, Umweltverschmutzung und Biodiversität versprechen, die sich in Realität aber nicht verbessern (OECD, 2025).

Technologien leisten in vielen Bereichen positive Beiträge, doch allein können sie unsere ökologischen und sozialen Probleme nicht lösen. Im Gegenteil sind sie oft selbst essentieller Teil dieser Krisen. Dieser Beitrag fokussiert auf Voraussetzungen und Wege, wie gesellschaftliche Akteure zusammenkommen können, um zu beginnen, die Richtung der technischen Entwicklungen aktiv zu steuern (McCormick et al. 2013) und die Rolle der Wissenschaften dabei. Nachhaltigkeit braucht richtungsweisende Entscheidungen, deren Bedeutung größere Aufmerksamkeit verdient, damit sowohl Entscheidungsträger:innen als auch Bürger:innen die Verantwortung für konkrete systemische Veränderungen bewusst übernehmen können. In ihrem Plädoyer für eine „Slow Science" schlägt Isabelle Stengers vor, dass Forschende beginnen sollten zuzuhören, auf die Weisheit des Kollektivs zu vertrauen und ihr Fachwissen einzusetzen, um die Anliegen der Öffentlichkeit zu unterstützen (Stengers, 2018). Nicht in erster Linie der wissenschaftliche Fortschritt oder wie die Wissenschaft alle Probleme der Gesellschaft lösen wird, sei wichtig.

Literatur

OECD (2025), Environmental Outlook on the Triple Planetary Crisis: Stakes, Evolution and Policy Linkages, OECD Environmental Outlook, OECD Publishing, Paris.

Stengers, I. 2018: Another Science Is Possible. A Manifesto for Slow Science. Cambridge, UK: Polity Press.

McCormick, K., Anderberg, S., Coenen, L., & Neij, L. (2013). Advancing sustainable urban transformation. Journal of Cleaner Production, 50, 1–11.

Sustainable AI and Energy Systems: Ethical and Epistemic Issues

Giovanni Frigo, Michael Poznic, Michael W. Schmidt (Karlsruhe Institute of Technology, Deutschland)

Sustainable AI, sustainable energy, ethics of technology, philosophy of technology, sociotechnical, energy systems

The integration of Artificial Intelligence (AI) in energy systems offers significant opportunities but also raises important ethical and epistemic challenges. This paper aims to connect the notion of "Sustainable AI" (van Wynsberghe 2021) and that of "Sustainable Energy" (e.g., Chu et al. 2017; Qazi et al. 2019; Halawa, 2024) to reflect on the adoption of AI in energy systems. We will address both (a) the energy-related sustainability of AI and (b) the use of AI for sustainability (van Wynsberghe 2021), that is, for achieving the goal of a comprehensively sustainable sociotechnical energy system (i.e., generation, distribution, use, waste, and recycling). The main ethical issue related to the sustainability of AI corresponds to its energy intensity (i.e., the energy consumption of large-scale AI systems). Ethical issues emerging from the use of AI for improving the sustainability of energy systems are, for example, fairness, responsibility, transparency, privacy, and (democratic) governance. Moreover, the deployment of AI for sustainability in energy systems presents significant epistemic challenges concerning, for example, knowledge production, uncertainty, and expertise. To illustrate these ethical and epistemic challenges in a concrete context, we examine the case of AI-based smart grid demand response systems. These systems use machine learning and predictive analytics to forecast electricity demand and automatically adjust consumption patterns, pricing signals, and distributed energy resources. As a prominent application of AI in contemporary energy systems, demand response provides a relevant case to explore issues of fairness, transparency, privacy, and uncertainty. Examining ethical and epistemic issues fosters clarifying value trade-offs, the criteria of responsible AI governance, and the design requirements of increasingly digitalized, decentralized, and renewable-based energy systems.