

## **Gemeinsam Erklär- und Kontrollierbarkeit schaffen: Partizipative Ansätze zur soziotechnischen Gestaltung des Einsatzes algorithmischer Systeme und künstlicher Intelligenz**

Annelie PENTENRIEDER, Ernst A. HARTMANN

*iiT – Institut für Innovation und Technik Berlin, in der VDI/VDE-IT GmbH  
Steinplatz 1, D-10623 Berlin*

**Kurzfassung:** Systeme, die auf fortgeschrittenen Algorithmen und großen Datenmengen basieren (auch als künstliche Intelligenz bezeichnet), kommen zunehmend in der industriellen Praxis zum Einsatz. Aufgrund ihrer Komplexität ist es jedoch häufig eine Herausforderung, sie gewinnbringend und risikoarm in bestehende Arbeitsabläufe einzubinden. Soziotechnische Perspektiven wie die nutzergerechte Erklär- und Kontrollierbarkeit für Fachkräfte mit unterschiedlichem Kenntnisstand werden dazu in Zukunft unverzichtbare Unterstützungen bieten. Im Vortrag berichten wir von einer Methode zur praktischen Erarbeitung von Erklär- und Kontrollierbarkeit komplexer Techniksysteme. Diese wurde in einem Workshop mit interdisziplinären Arbeitsgruppen erprobt. Zur Ergebnissicherung wurden grafische Visualisierungen eingesetzt. Ziel ist es, praxisorientierte Vorgehensweisen zur Gestaltung von Erklär- und Kontrollierbarkeit zu finden, die langfristig auch eine soziotechnische Überprüfung im Nutzungskontext im Rahmen aktuell geplanter Zertifizierungen und Auditierungen von KI-Systemen sicherstellen können.

**Schlüsselwörter:** Erklärbare künstliche Intelligenz, partizipative Technikgestaltung, nutzerzentriertes Technikdesign, Co-Creation-Workshops, nutzerzentrierte Visualisierung, Fachkräfte

### **1. Soziotechnische Perspektiven auf künstlich-intelligente Systeme etablieren**

Wer heute in der Arbeitswelt eine Software nutzt, wurde überwiegend nicht an deren Entwicklung beteiligt. Die Software ist häufig nicht auf die individuellen Arbeitspraktiken abgestimmt und erfordert von den Anwendenden viel Geduld und technischen Sachverstand in der Anpassung an die erforderlichen Aufgaben im Arbeitsalltag. Standardisierte Softwaresysteme werden häufig fernab vom tatsächlichen Einsatz im Alltagskontext entwickelt, ihre Entwickler:innen verwenden dazu teilweise „Ich-Methodologien“: Die französische Techniksoziologin Madeleine Akrich bezeichnet die Methode, über eigene Erfahrungen und Situationsbewertungen zu einer Nutzerrepräsentation zu gelangen, als „i-methodology“. Es handelt sich um einen unbewussten Prozess, in dem auf die persönliche Erfahrung vertraut wird und der/die Gestalter:in seine/ihre professionelle Rolle – „his professional hat“ – mit dem des Laien ersetzt. Solche „impliziten“ Methoden sind in der Softwareentwicklung häufig mächtiger als „explizite“ Nutzerforschungen, die Nutzende womöglich erwarten würden (Akrich 1995). Somit nimmt die Software die Gestalt an, die für die Entwickler:innen selbst am praktischsten wäre. Gerade die Bedarfe von marginalisierten Nutzergruppen (z.B. Frauen in technischen Berufen, Werker:innen mit

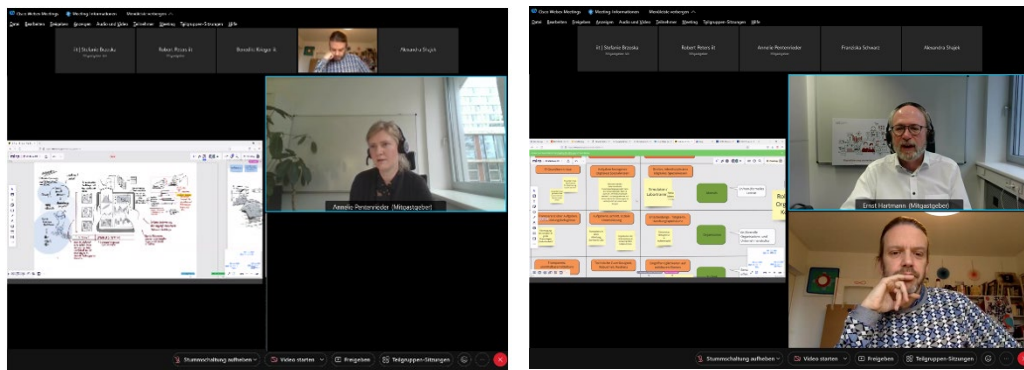
Beeinträchtigungen, individualisierte Bedarfe in Produktionsbetrieben mit Spezialanfertigungen) werden dabei zu wenig berücksichtigt.

Insbesondere in der europäischen KI-Forschung werden aktuell erklärbare und transparente KI-Systeme diskutiert, damit komplexe Systeme in Zukunft auch von einer Vielzahl an Nutzer:innen erschlossen und verstanden werden können. Für partizipative Ansätze bietet dieser Bewusstseinswechsel eine neue Chance. Selbst Softwareunternehmen bestätigen inzwischen den Mehrwert, diverse Nutzergruppen an Technikgestaltung zu beteiligen. Rosmarie Steininger vom Softwareunternehmen Chemistree geht sogar einen Schritt weiter und berichtet, dass die relevanten Fragen zur Softwaregestaltung häufig von Personen ohne IT-Background gestellt werden (Steininger 2021). Im Rahmen der Initiative KIDD — KI im Dienste der Diversität — plädiert sie darum für den Ansatz „Innovation by Diversity“ (KIDD o.J.).

Die Gestaltung von Formaten für diese vielfältige Einbindung ist eine Herausforderung an sich. An drei Use Cases der Universität Bamberg, der Universität Stuttgart und der Universität Dortmund wurde in einem Workshop des Instituts für Innovation und Technik Berlin (iit) im Dezember 2021 erprobt, wie KI-basierte Techniksysteme erklär- und kontrollierbar gemacht werden können (iit 2021). Dabei wurden insbesondere auch Zielgruppen wie Frauen oder Nicht-Techniker:innen als Nutzende bedacht, um vielfältigere Zugänge zu komplexen Systemen zu finden. Um niedrigschwellige Partizipationsansätze methodisch zu erproben, wurden in den Workshops visuelle Dokumentationsformate eingebunden. Daran wurde konkret diskutiert, an welchen Stellen in der Softwaregestaltung welcher Einfluss von Anwendenden einer Software produktiv ist.

## **2. Partizipatives Design methodisch neu gedacht in Co-Creation-Workshops mit Live-Visualisierung**

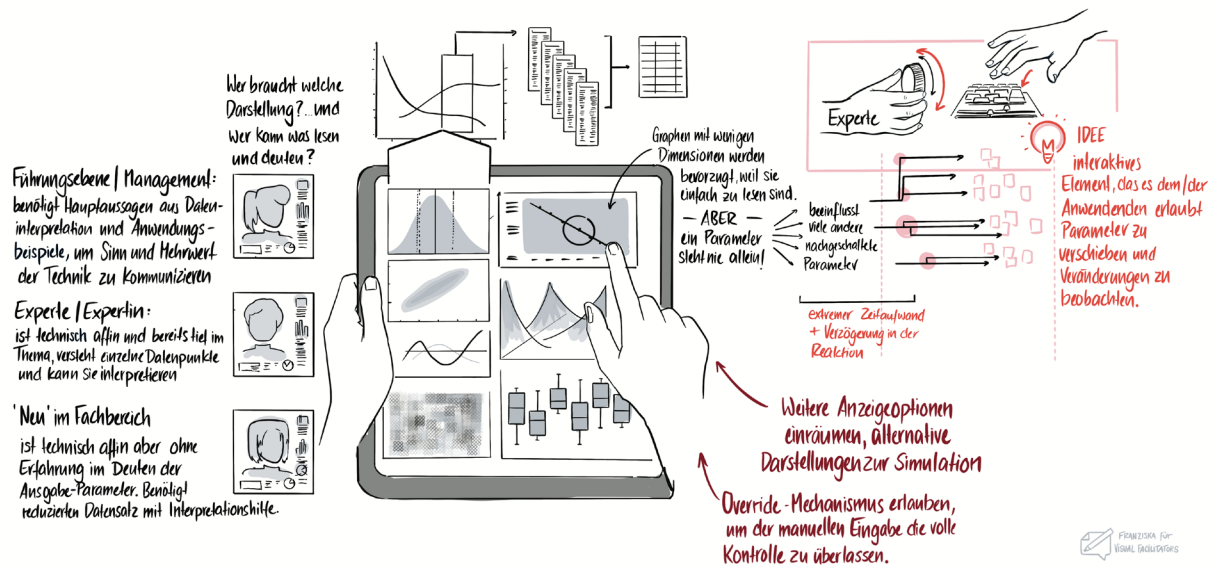
Diverse Anwender:innen an Technikgestaltung zu beteiligen hat eine lange Tradition. In der skandinavischen Forschungsrichtung des „Participatory Design“ beschäftigen sich Technikdesigner:innen seit den frühen 70er Jahren mit der Einbindung von Nutzer:innen in die Designprozesse ihrer eigenen Technologien (Ehn 2010; Greenbaum & Kyng 1991; Simonsen & Robertson 2013). Bereits damals entstand durch die Entwicklung neuer Computertechnologien die Sorge, die neuen Technologien könnten zur Dequalifizierung von Arbeitskräften beitragen. Im technikintensiven Arbeitskontext bildeten sich darum Teams aus Angestellten und Gewerkschaften mit dem Ziel, Technik mit Rücksicht auf Teilhabe am Arbeitsplatz so zu gestalten, dass sie Arbeiter:innen und Unternehmer:innen gleichermaßen dient. Technologien sollten es den Anwender:innen ermöglichen, eigenes implizites Wissen und spezifische Kompetenzen mit automatisierten Prozessen zusammenzubringen. Im „Participatory Design“ wurden darum methodische Ansätze entwickelt, um Technologien nicht nur für sondern auch mit den Nutzer:innen zu gestalten. Zwar entwickelt diese Forschungstradition bis heute alternative Innovationskonzepte, die kooperative Technologien hervorbringen. Jedoch war diese Art der Technikgestaltung auch im politischen Zeitgeist der 70er/80er Jahre beheimatet, in der etwa Gewerkschaften und andere Arbeitnehmervertreter auch in Deutschland eine viel präzisere Rolle als heute in der Technikgestaltung hatten (Napper 1994). Es gilt, diese Ansätze mit den digitalen Möglichkeiten der heutigen Zeit gerade im Kontext der Nutzung von KI-Systemen wieder stark zu machen.



**Abbildung 1:** © iit Berlin. Erarbeitung der Visualisierungen zu Erklär- und Kontrollierbarkeit künstlich intelligenter Systeme mit dem Konferenztool Webex und der Kollaborationsplattform Miro.

In einem offenen Workshop mit dem Titel „So sehen sie aus – erklär- und kontrollierbare IT-Systeme“ wurde entlang von Use Cases aus dem Automobilwesen, aus dem Brauereiwesen und aus der Bereich des Anlagenbaus für Abfüllanlagen Software zur Einbettung in industrielle Arbeitsprozesse betrachtet. Im Projekt der Universität Bamberg und der Continental AG werden beispielsweise Ingenieur:innen bei der Entwicklung eines Autos unterstützt: Es steht eine Steuerungssoftware im Zentrum, die für den Kofferraumdeckel notwendig ist, um Absenkgeschwindigkeit und -kraft in Abhängigkeit verschiedener Parameter zu steuern und Einklemmen zu vermeiden. An diesem Beispiel wird von der Universität Bamberg gemeinsam mit dem Fraunhofer Institut erforscht, wie lernende Systeme und menschliche Domänenexpert:innen wechselseitig voneinander profitieren können. In ähnlicher Weise geht es im Projekt der TU Dortmund um die maschinelle Analyse von Sensordaten aus dem Brauprozess mit dem Ziel, bestimmte Aspekte des Biergeschmacks gezielt beeinflussen zu können. Die Universität Stuttgart beschäftigt sich wiederum in ihrem Projekt mit der Unterstützung von Arbeitenden bei der Umrüstung von Abfüllanlagen mit Hilfe von Sensorik und KI.

Ziel des Workshops war es, Beispielszenarios zu erarbeiten, in denen Erklär- und Kontrollierbarkeit von IT-Systemen konkret gestaltet wird. Der Ablauf war von zwei Besonderheiten gekennzeichnet. Zum einen wurde die Partizipation der Teilnehmenden dadurch ermöglicht, dass die virtuelle Online Kollaborations-Plattform für Teams „Miro“ genutzt wurde. Dieses Whiteboard diente dazu, eine Vielfalt an Expert:innenstimmen einzubinden und im späteren Verlauf Ideen und Erfahrungen ordnen und rekonstruieren zu können. Die Diskussion wurde durch eine Beschreibungsmatrix strukturiert, die eine soziotechnische Sicht auf die Systemgestaltung abbildet. In den drei Zeilen der Matrix finden sich die Teilsysteme eines soziotechnischen Systems – Technik, Organisation und Mensch. Die drei Spalten bilden Aspekte ab, die die Erklär- und Kontrollierbarkeit des KI-geprägten Arbeitssystems aus handlungs- und kontrolltheoretischer Sicht beschreiben: Transparenz, Handlungssicherheit (Effizienz) und Handlungsspielraum (Divergenz). Zu allen neun Aspekten, die sich aus den Zellen der Matrix ergeben, wurden von den am Co-Creation-Workshop Beteiligten spezifische Merkmale herausgearbeitet. Diese Aspekte können weiterentwickelt werden zu einem Instrument der Auditierung von KI-Anwendungsumgebungen (Hartmann 2021, dort auch weitere Literaturangaben).



**Abbildung 2:** © iit Berlin/Visual Facilitators. Ergebnisansicht eines Co-Creation-Workshops. Entlang von Mockups für erklärbare-KI-Benutzeroberflächen haben Teams die Möglichkeit, verschiedene Szenarien für ihren Arbeitskontext zu erproben. Gerade für partizipative Workshops eignet sich die Einbindung von Visualisierer:innen, da sie eine niedrigschwellige Zugänglichkeit ggf. auch für nicht-technisch-affine Personen und eine direkte und spielerische Anpassung über die Mockups möglich ist.

Für die Herausforderung, dass Erklär- und Kontrollierbarkeit von KI häufig nur schwer konkretisiert werden kann, wurde zudem ein zweites Vorgehen erprobt: Für jeden Beispielfall wurden Zeichner:innen in die Diskussion eingebunden, die die diskutierten Szenen und Beispielszenarios live grafisch darstellten. Die Visualisierungen konkretisierten Ideen und weitere Diskussionsbedarfe zur Erklär- und Kontrollierbarkeit von Systemen, etwa die Darstellung interaktiver Benutzerschnittstellen (visualisiert in Mock-ups) und neuartige Visualisierungsmethoden für die Nutzer:innen zu wissenswerten technischen Prozessen im Hintergrund. Des Weiteren machten die Visualisierungen sichtbar, welche analogen Handarbeiten der Einsatz von komplexen IT-Systemen nach wie vor erfordert, die im Team gut moderiert und sozial begleitet werden müssen.

### 3. Ergebnis und Ausblick

Partizipative Workshop-Formate können das Zusammenbringen unterschiedlicher Kompetenzen zur Bedarfsabfrage für die Technikgestaltung erleichtern und eine Einbindung verschiedener Interessengruppen in die Technikgestaltung ermöglichen. Bei allen drei Fallbeispielen war es das Ziel, erklär- und kontrollierbare Systeme zu schaffen, die Expert:innen oder angehende Expert:innen einer bestimmten Domäne dabei unterstützen, (kausale) Beziehungen in komplexen Daten zu identifizieren. Zudem zielen erklär- und kontrollierbare Systeme darauf ab, dass versierte Fachkräfte ihr Erfahrungswissen in den maschinellen Lernprozess moderner KI-Systeme einbringen und an weniger erfahrene Kolleg:innen weitergeben.

Der Workshop hat ergeben, dass die Beteiligung praktischer Anwender:innen bereits in der Akquise gut organisiert werden muss. Zudem muss geklärt werden, aus welcher Nutzersicht die Software gestaltet werden sollte (hier sind mehrere aufeinander aufbauende Workshops hilfreich). Auch mithilfe von Personas

(stereotypisierte Anwender) kann zwischen individueller Anpassung und Standardangeboten einer Software abgewogen werden:



**Abbildung 3:** © iit Berlin. Darstellung von Personas. Software-Ingenieur, Produktionsleiterin oder Werkzeugmacher haben unterschiedliche Anforderungen an eine Software und deren Handhabung und Einbindung im Alltag. Ihre Anliegen sollten von Anfang an in die Softwaregestaltung miteinbezogen werden. Partizipatives Technikdesign ermöglicht es auch, die Gestaltung der Technik von verschiedenen Perspektiven im Arbeitsprozess aus kontinuierlich zu überprüfen.

Gerade im Maschinenbau kann die Öffnung der Technikgestaltung für eine diverse Anwenderschaft zwei zukünftigen Herausforderungen Rechnung tragen. Zum einen wirkt eine diverser gestaltete Partizipation dem Fachkräftemangel entgegen (z.B. mehr Frauen in der Industrie 4.0), zum anderen erfordert die in Zukunft von Unternehmen geforderte Informationsfähigkeit über CO<sub>2</sub>-Einsparungen, dass Daten von unterschiedlichen Fachkontexten aus überprüft werden können. Einer heterogenen, interdisziplinären Nutzerschaft sollte demnach in Zukunft sehr viel mehr Einsicht in technische Abläufe und Daten ermöglicht werden. Eine Einbindung diverser Gruppen muss bereits früh im Prozess mitgedacht werden, denn bei der Wahl von Daten, Sensoren, Algorithmen und Softwarearchitektur (z.B. deterministisches System oder lernendes System) handelt es sich häufig um Grundsatzentscheidungen, die später nur schwerlich revidiert werden können.

Soziotechnische Aspekte wie die Erklär- und Kontrollierbarkeit komplexer Systeme im Kontext von Technik, Organisation und Menschen rücken zunehmend auch in den Fokus der politischen Agenda (DIN/DKE 2020, dort insbes. Arbeitsgruppe 3). Der AI Act der Europäischen Kommission berücksichtigt die Beteiligung heterogener Anwender:innen mit unterschiedlichem Kenntnisstand, indem die Notwendigkeit menschlicher Aufsicht und spezielle Dokumentationspflichten von KI-Systemen gefordert werden (EU 2021). Technische Systeme, die diese heterogene Nutzung bereits von Beginn der Technikgestaltung an eingepreist haben, werden hier im Vorteil sein.

#### 4. Literatur

- Akrich M (1995) User Representations: Practices, Methods and Sociology. Managing Technology in Society. In: Rip A, Schot J, Misa T J (Hrsg) The Approach of Constructive Technology Assessment. London / New York: Pinter Publisher, 167–184.
- DIN/DKE Deutsches Institut für Normung, Deutsche Kommission für Elektrotechnik (2020) Deutsche Normungsroadmap Künstliche Intelligenz, Berlin, Frankfurt am Main. Abgerufen am 12. Dezember, 2021. [www.din.de/resource/blob/772438/6b5ac6680543eff9fe372603514be3e6/normungsroadmap-ki-data.pdf](http://www.din.de/resource/blob/772438/6b5ac6680543eff9fe372603514be3e6/normungsroadmap-ki-data.pdf).
- Ehn P (2010) Exploring the Scandinavian Participatory Design Tradition. Medea Talks #17, School of Arts and Communication, Malmö. <https://medium.com/@medea/design-democracy-and-participation-exploring-the-scandinavian-participatory-design-tradition-67dda92ba675>.
- EU – Europäische Union (2021) Proposal for a Regulation of the European Parliament and of the Council laying down harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts. Abgerufen am 12. Dezember, 2021. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206>.
- Greenbaum J, Kyng M (1991) Design at Work: Cooperative Design of Computer Systems, Hillsdale, N.J: Erlbaum.
- Hartmann E A (2021) Digitale Souveränität: Soziotechnische Bewertung und Gestaltung von Anwendungen algorithmischer Systeme. In: Hartmann E A (Hrsg.) Digitalisierung souverän gestalten II – Handlungsspielräume in digitalen Wertschöpfungsnetzwerken. Wiesbaden: Springer Vieweg.
- iit – Institut für Innovation und Technik (2021) Symposium Digitalisierung souverän gestalten II. Abgerufen am 12. Dezember, 2021. [www.iit-berlin.de/symposium-digitalisierung-souveraen-gestalten-ii/](http://www.iit-berlin.de/symposium-digitalisierung-souveraen-gestalten-ii/).
- KIDD (o.J.): KIDD – Künstliche Intelligenz im Dienste der Diversität [www.kidd-prozess.de](http://www.kidd-prozess.de). Abgerufen am 12. Dezember, 2021.
- Napper V S (1994) A Review: Design at Work: Cooperative Design of Computer Systems. Educational Technology Research and Development 42.1:97–99.
- Simonsen J, Robertson T (2013) Routledge International Handbook of Participatory Design, New York: Routledge.
- Steininger R (2021) Vortrag beim ersten Round-Table der Reihe „Digitalisierung geschlechtergerecht gestalten“. Abgerufen am 16. Dezember, 2021. <https://www.eaf-berlin.de/news/digitalisierung-geschlechtergerecht-gestalten/>.

**Danksagung:** Ein besonderer Dank gilt den Teilnehmerinnen und Teilnehmern des Workshops „So sehen sie aus – erklär- und kontrollierbare IT-Systeme“ im Rahmen des II. Symposiums „Digitalisierung souverän gestalten“ des Institut für Innovation und Technik Berlin am 1. Dezember 2021 im Auftrag der Heidenhain GmbH.



Gesellschaft für  
Arbeitswissenschaft e.V.

## Technologie und Bildung in hybriden Arbeitswelten

68. Kongress der  
Gesellschaft für Arbeitswissenschaft

Otto-von-Guericke-Universität Magdeburg

Fraunhofer-Institut für Fabrikbetrieb und  
Fabrikautomatisierung IFF, Magdeburg

02. – 04. März 2022

---

## GfA-Press

---

**Bericht zum 68. Arbeitswissenschaftlichen Kongress vom 02. – 04. März 2022**

**Otto-von Guericke-Universität Magdeburg;  
Fraunhofer-Institut für Fabrikbetrieb und -automatisierung IFF, Magdeburg**

Herausgegeben von der Gesellschaft für Arbeitswissenschaft e.V.  
Sankt Augustin: GfA-Press, 2022  
ISBN 978-3-936804-31-7

NE: Gesellschaft für Arbeitswissenschaft: Jahresdokumentation

Als Manuskript zusammengestellt. Diese Jahresdokumentation ist nur in der Geschäftsstelle (s. u.) erhältlich.

Alle Rechte vorbehalten.

© **GfA-Press, Sankt Augustin**

**Schriftleitung: Prof. Dr. Rolf Ellegast**

im Auftrag der Gesellschaft für Arbeitswissenschaft e.V.

Ohne ausdrückliche Genehmigung der Gesellschaft für Arbeitswissenschaft e.V. ist es nicht gestattet:

- den Kongressband oder Teile daraus in irgendeiner Form (durch Fotokopie, Mikrofilm oder ein anderes Verfahren) zu vervielfältigen,
- den Kongressband oder Teile daraus in Print- und/oder Nonprint-Medien (Webseiten, Blog, Social Media) zu verbreiten.

Die Verantwortung für die Inhalte der Beiträge tragen alleine die jeweiligen Verfasser; die GfA haftet nicht für die weitere Verwendung der darin enthaltenen Angaben.

### **Geschäftsstelle der GfA**

Simone John, Tel.: +49 (0)30 1300-13003

Alte Heerstraße 111, D-53757 Sankt Augustin

[info@gesellschaft-fuer-arbeitswissenschaft.de](mailto:info@gesellschaft-fuer-arbeitswissenschaft.de) · [www.gesellschaft-fuer-arbeitswissenschaft.de](http://www.gesellschaft-fuer-arbeitswissenschaft.de)

### **Screen design und Umsetzung**

© 2022 fröse multimedia, Frank Fröse

[office@internetkundenservice.de](mailto:office@internetkundenservice.de) · [www.internetkundenservice.de](http://www.internetkundenservice.de)