

## **Hybride Empirie – Digitale Verhaltensdaten als WIRKSames Analyseinstrument der Interaktion mit Künstlicher Intelligenz**

Andrea ALTEPOST, Adjan HANSEN-AMPAH, Wolfgang MERX

*Institut für Textiltechnik, RWTH Aachen University  
Otto-Blumenthal-Str. 1, D-52074 Aachen*

### **Kurzfassung:**

Das Kompetenzzentrum für Arbeitsforschung WIRKSam erforscht seit 01.11.2021 arbeitswissenschaftliche Aspekte des Einsatzes Künstlicher Intelligenz in der Textil- und Metallbranche im Rheinischen Braunkohlerevier. Damit leistet es einen Beitrag zum Strukturwandel vor dem Hintergrund des Kohleausstiegs. Für zunächst neun (und ab 2024 weitere sechs) Anwendungsfälle werden in regionalen Unternehmen partizipativ KI-Anwendungen entwickelt, arbeitswissenschaftlich gestaltet und analysiert. Eine der Komplexität des Projekts angemessene Methodik der empirischen Analyse muss sowohl übergreifende als auch fallspezifische Zielstellungen adressieren. Zur empirischen Untersuchung und Evaluation der betrachteten soziotechnischen Systeme werden daher unter anderem Digitale Verhaltensdaten genutzt, die aus der Interaktion der Beschäftigten mit den KI-Systemen emergieren. Neben der jeweils spezifischen Evaluation der Anwendungen im Rahmen einer iterativen, partizipativen Entwicklung und Interpretation der Daten im Kontext des jeweiligen Unternehmensbeispiels können übergreifende Aussagen etwa zur Handhabung und Nutzung der KI oder Belastungsfaktoren abgeleitet werden. Neben Logfiles können hier z. B. auch Bewegungsdaten herangezogen und integriert mit „klassischen“ Befragungs- und Beobachtungsdaten ausgewertet werden.

**Schlüsselwörter:** Kompetenzzentrum, Strukturwandel, Digitale Verhaltensdaten, Künstliche Intelligenz, Soziotechnisches System, Partizipation

## **1. Einleitung**

Digitale Verhaltensdaten (DVD) werden derzeit von Expert\*innen sozialwissenschaftlicher Methoden verstärkt in den Blick genommen, so auch bei der GESIS (vgl. GESIS – Leibniz-Institut für Sozialwissenschaften, o. D.). Im Fokus stehen meist Daten aus sozialen Netzwerken oder Geodaten, die unabhängig von der empirischen Untersuchung entstehen. Die Arbeitswissenschaft andererseits diskutiert die Adaption bzw. Entwicklung geeigneter Instrumente zur empirischen Analyse soziotechnischer Systeme im digitalen Kontext. Die komplexe Interaktion zwischen Mensch und digitaler Technik stellt – dem Techniksoziologen Häußling folgend – im Kern eine Datenverkopplung dar (Häußling et al. 2017) und bedarf zu ihrer vollständigen Analyse auch der Betrachtung der technisch erzeugten Daten im „Handshake“ zwischen Mitarbeiter\*in und dem digitalen System. Während im GESIS-Ansatz digitale Verhaltensdaten mit Survey-Daten kombiniert werden, um die Vor- und Nachteile der unterschiedlichen Methoden optimal auszubalancieren, geht es darüber hinaus in WIRKSam vor allem darum, die Arbeitssituation mit der zentralen Interaktion zwischen Mensch und Künstlicher Intelligenz (KI) ganzheitlich zu erfassen. Daher ist geplant, weitere – auch

technisch erfassbare – Daten wie etwa Bewegungs- und Körperdaten von Nutzer\*innen im Rahmen der anstehenden Nutzertests ebenso wie „klassische“ Daten aus Beobachtung und Befragung zu erheben. Dieses Setting wird in Nutzertests der aktuell neun zu entwickelnden Demonstratoren angewendet werden und ist in zweierlei Hinsicht zu differenzieren:

- **Evaluation für die Technikentwicklung vs. Erkenntnisse zur Gestaltung des Arbeitssystems/Qualifizierung**  
Die Daten sollen zur Evaluation prototypischer Entwicklungsstufen der KI beitragen und für weitere Iterationen Empfehlungen liefern; außerdem sollen die Bedingungen der Nutzung durch Mitarbeitende erforscht und Belastungspotenzial identifiziert werden, um entsprechende Zusammenhänge im Kontext des MTO-Prinzips (simultane Betrachtung von Mensch, Technik und Organisation; Strohm und Ulich 1997) aufzudecken und Empfehlungen für die Arbeitsgestaltung einschließlich Qualifizierungsbedarf zu generieren.
- **Unternehmensspezifischer Anwendungsfall im Vergleich zur Gesamtheit der Erfahrungen mit den WIRKSAM-Demonstratoren**  
Die Ergebnisse der Evaluation und der Arbeitsgestaltungsanalyse im einzelnen Use Case sind in Beziehung zu setzen zu den Resultaten der anderen Demonstratoren. Es muss überprüft werden, welche Aussagen über Bedingungen, Determinanten (z. B. sozioökonomische Faktoren), erfolgversprechende Vorgehensweisen und Auswirkungen der Arbeitsgestaltung mit KI getroffen werden können und welche Rolle die Spezifika des jeweiligen Unternehmens und seiner Anwendung spielen.

## 2. Methodik

Der geplante Ansatz basiert auf dem technografischen Forschungsprogramm von Rammert (vgl. Rammert 2007). Er fokussiert damit das verteilte Handeln zwischen Mensch und Technik und rekurriert auf eine ethnografische (qualitative) Grundlage in Kombination mit der Verwendung auch technisch erzeugter Daten. Diese legen eine quantitative Auswertung nahe, sodass ein umfassender Mixed-Methods-Ansatz entsteht. Seinen Namen verdient dieser allerdings erst dann, wenn die Daten verschiedener Provenienz in integrierter Weise ausgewertet werden. Die Leitplanken des Auswertungskonzepts werden sich daher zunächst an dem folgenden Vorgehen orientieren (Altepost et al. 2021, S.84):

- **Quantitative Auswertung unter Berücksichtigung qualitativer Daten** (z. B. Zusammenhang von Technikaffinität mit im Interview berichteter Freude an der Interaktion mit der KI)
- **Qualitative Auswertung unter Berücksichtigung quantitativer Daten**: z. B. Differenzierung beobachteter Merkmale nach soziodemographischen Aspekten wie Alter oder Bildungsniveau)
- **Auswertung von Logfile-Daten zur Analyse der „verkoppelten“ Vorgänge in der Mensch-Technik-Interaktion**
- **Interpretierende Gesamtschau**: Theoriegeleitete Interpretationen sind ebenso möglich wie die explorative Generierung von Hypothesen. Beispielsweise können die Logfiles in Verbindung mit quantitativen Daten (z. B. Zusammenhang der Häufigkeit von

fehlerhafter Interaktionen mit Stress im NASA-TLX) sowie qualitativen Daten (z. B. Beobachtung besonderer Ereignisse wie mehrfache Ausführung desselben Tätigkeitsschritts und möglicher Zusammenhang mit unintendierten Aspekten der Menünutzung) gebracht werden. Von Interesse wäre z. B. im Sinne einer ganzheitlichen Betrachtung, welche Zustände oder Ereignisse der Gesamtsituation mit Auffälligkeiten in den Logdateien in einen Zusammenhang zu bringen sind und wie diese Koinzidenz zu interpretieren ist (z. B. in einem Sinne von Kausalität).

### 3. Erste konzeptuelle Schritte

Um die Basis zu einem übergreifenden Konzept zur Identifizierung und Gewinnung relevanter technischer Daten in Tests prototypischer KI-Versionen zu legen, wurden zunächst Anforderungen an digitale Verhaltensdaten aus den Zielen der empirischen Analyse abgeleitet. Diese befinden sich zum derzeitigen Stand des Projekts noch auf einem relativ allgemeinen Level, was auch mit der partizipativen Ausrichtung der Vorgehensweise zu tun hat. In Anlehnung an die APRODI-Studie (Schmidt et al. 2020) und weitere Modelle für Digitalisierungsprojekte (SozioTex-Vorgehensmodell / Altep post et al. 2021; Plattform Industrie 4.0 / Stowasser et al. 2020; VDI-VDE 3714, 2020) wird derzeit ein Vorgehensmodell erarbeitet, das die oben skizzierte Dialektik der vergleichenden, übergreifenden Ableitung von Ergebnissen des Gesamtprojekts mit der Berücksichtigung individueller Spezifika der Use Cases verbinden soll. Die methodisch-empirische Vorgehensweise ist ein Aspekt davon. Kriterien der Betrachtung werden entsprechend teils im Konsortium standardisiert, teils im Team des jeweiligen Use Cases mit den Unternehmen und ihren Mitarbeitenden erarbeitet. Dennoch lassen sich einige grundlegende Eckpunkte der Verwendung digitaler Verhaltensdaten benennen.

#### 3.1 Anforderungen an digitale Verhaltensdaten

Aus den Zielen der empirischen Analyse ergeben sich für das Analysekonzept einige Konsequenzen für die Erhebung digitaler Verhaltensdaten. So muss die Erhebung von Screenrecordings, Logfiles und gegebenenfalls Eyetracking-Daten übergreifende Features beinhalten, z. B.:

- Zeitstempel/Verweilen auf Seiten und Reihenfolge, in der sie aufgesucht werden – wie bewegt sich der/die Nutzer/in durch die Anwendung?
- Augenbewegung – Gibt es Anzeichen einer zu komplexen Benutzeroberfläche oder gar Ablenkung, z. B. durch Animationen?
- Klicken von Buttons (zusätzliche Informationen über Seitenwechsel hinaus, z. B. Vergrößerung von Darstellungen o. ä.; werden auf einer Seite eingebettete Erläuterungsvideos verwendet etc.); – hiermit kann z. B. untersucht werden, wonach sich die Anzahl der Seitenwechsel richtet (z. B. sehr hohe Anzahl: sachlich begründet oder evtl. Hinweis auf Qualifikationsbedarf oder nicht-intuitive Menüführung?)
- Fehler in der Interaktion mit der KI; Hinweis auf Qualifikationsbedarf oder nicht-intuitive Menüführung?

Um die digitalen Verhaltensdaten für die interpretierende Gesamtschau anschlussfähig zu machen, müssen die Zeitstempel mit den Beobachtungen äußeren Handelns und relevanter Ereignisse (z. B. Zeitstempel im Videomaterial) synchronisiert werden.

### *3.2 Entwicklung neuer Indikatoren für die Erfassung spezifischer Eigenschaften von KI*

Hier gilt es, beispielsweise die Wirkung erklärbarer KI („Explainable AI“) zu messen. Was kann in KI-Anwendungen überhaupt dem Nutzer/der Nutzerin transparent gemacht werden? Wie ist hierbei der Erfolg messbar? Ein möglicher Ansatz zur Beantwortung dieser Fragen ist die Erfassung der Akzeptanz von Nutzer\*innen in drei Anwendungsszenarien. Im ersten Szenario wird ohne KI gearbeitet. Im zweiten Szenario kommt eine KI zum Einsatz, die wie eine Blackbox fungiert und für Nutzer\*innen nicht ohne weiteres nachvollziehbare Lösungen generiert. Im dritten Szenario werden mithilfe von Explainable AI Ergebnisse produziert, die von Nutzer\*innen im besten Falle als nachvollziehbar wahrgenommen werden. Auf Grundlage der in diesen drei Szenarien jeweils generierten Akzeptanz können Rückschlüsse auf den Erfolg der KI-Anwendungen gezogen werden. Eine dazu letztlich komplementäre Möglichkeit wäre, die mentalen Modelle der Nutzer\*innen über die Arbeitsweise einer KI zu analysieren und auf ihr Zutreffen hin zu untersuchen. Ob und inwiefern die Nachvollziehbarkeit der KI sich auf Vertrauensbildung und Akzeptanz auswirkt, wäre ebenfalls empirisch zu prüfen.

### *3.3 Validierung*

Die integrierende Gesamtschau hat eine Kehrseite, die auf eine Reihe methodischer Problemstellungen aufmerksam macht. So ist hier erneut zu hinterfragen, welche methodischen Konstrukte mit den jeweiligen Instrumenten gemessen bzw. betrachtet werden und an welchen Kriterien diese jeweils bislang validiert wurden. Beispielsweise wurde (in einer als eher explorativ zu beurteilenden Studie) festgestellt, dass Aussagen zu Stressempfinden im NASA-TLX nicht mit beobachteten Stressreaktionen korrelierten (Altepost et al. 2021). Die jeweiligen Indikatoren müssen also daraufhin untersucht werden, welches Verständnis von Stress genau die Befragten im NASA-TLX ihrer Aussage zugrunde legen (z. B. Zeitdruck, Empfinden der Mensch-Technik-Interaktion als unangenehm, Überforderung durch die mit der KI zu lösenden Aufgabe) und wie Indikatoren beispielsweise in der Beobachtung angewendet werden (welche Verhaltensweisen deuten beispielsweise valide auf eine Stresssituation hin?). Gegebenenfalls sind etablierte arbeitswissenschaftliche Instrumente durch andere Inventare zur Messung bestimmter Schwerpunkte, etwa aus dem Stresserleben, zu ergänzen.

### *3.4 Verdichtung der technischen und nicht-technischen Indikatoren zu einem validen und praktikablen Verfahren zur Analyse von Nutzertests*

Die iterative Evaluation von KI und die simultane Arbeitsgestaltung sind mit hohem Zeit- und Personalressourcen-Aufwand verbunden. Mag dies im Rahmen eines Forschungsvorhabens noch realisierbar sein, ergeben sich für eine innerbetriebliche partizipative Innovationsgestaltung jedoch beträchtliche Hürden. Ein Ziel ist, diesen Aufwand möglichst weit zu reduzieren, sodass die resultierende Verfahrensweise weiterhin valide Ergebnisse und damit verlässliche Handlungsempfehlungen erbringt. Ganz ohne Aufwand – z. B. Workshops mit den betrieblichen Stakeholdern, umfängliche Nutzertests – wird es allerdings nicht gehen. Hier könnte eine Abschätzung der Benefits partizipativer Gestaltung nach dem MTO-Prinzip für die Wirtschaftlichkeit und Qualität der Arbeit mit KI helfen.

### **3.5 Beispiel: Analyse eines Demonstrators zur KI-Qualifizierung für qualitätskritische Fertigkeiten**

Beispielhaft sei ein Anwendungsfall betrachtet, in dem ein interaktiver, KI-gestützter Roboter entwickelt werden soll, der in der Lage ist, motorische Arbeitsstrategien erfahrener Beschäftigter zu erfassen, zu analysieren und an weniger erfahrene Beschäftigte weiterzugeben. Die zu erlernenden komplexen motorischen Tätigkeiten dienen der Herstellung von Faserverbund-Bauteilen, z. B. für den Automobilbau. Eine spezifische Ausbildung gibt es hierfür bislang nicht. Bereits kleine Fehler in der Ausführung können ein Bauteil für die vorgesehene Verwendung unbrauchbar machen.

Neben der Logfile-Aufzeichnung der Interaktion zwischen Mensch und KI, die Aufschluss über die Anwendung der KI-App beim Lernen am Arbeitsplatz gibt, können zudem Bewegungsdaten über die Verknüpfung mit Qualitätsdaten den Mitarbeitenden Feedback geben. Insgesamt werden Erkenntnisse über erfolgreiche und weniger erfolgreiche motorische Strategien, aber auch über den Prozess des motorischen Lernens gewonnen. Aus der MTO-Perspektive kann damit beispielsweise auch die Rolle des/der betrieblichen Ausbilders/Ausbilderin neu gestaltet werden.

## **4. Diskussion, Fazit und Ausblick**

Die Analyse digitaler Verhaltensdaten im Projekt WIRKsam bietet die Chance, die Vorgänge in der Mensch-Technik-Interaktion am Arbeitsplatz bzw. beim Lernen in der Arbeit differenziert zu beschreiben und besser zu verstehen. Dies erfordert auch die Klärung methodischer Fragen, etwa nach der Validität der technischen und auch der nicht-technisch generierten Daten im integrierten Kontext. Für die Methodenentwicklung der empirischen Arbeitsforschung steckt hierin großes Potenzial. Gerade die Komplexität in WIRKsam mit seinen derzeit neun und insgesamt 15 geplanten Anwendungsfällen ermöglicht auch fallabstrahierende Erkenntnisse über die Arbeitsgestaltung mit KI und insbesondere die Mensch-Technik-Interaktion. Die Spezifikation der Use Cases erfolgt partizipativ in den Unternehmen, wobei auch die konkret benötigten Daten und deren Zielsetzung in der Analyse festgelegt und jeweils ein teils standardisiertes, teils individuelles Design erarbeitet werden muss. Hinsichtlich des Datenschutzes ist hier neben Privacy (anonymisierte Erhebung, die Rohdaten dürfen nicht dem jeweiligen Unternehmen zugänglich gemacht werden), Safety und Security auch der Aspekt der Datensparsamkeit zu nennen. Die hier im Rahmen der Forschung gesetzten Standards sind nicht ohne weiteres auf die spätere betriebliche Nutzung zu übertragen: Sollen in der betrieblichen Anwendung weiterhin Daten getrackt werden – z. B. zu Lernzwecken wie im oben geschilderten Beispiel –, muss dies mit den Mitarbeitenden verhandelt und vertraglich bzw. durch Betriebsvereinbarung geregelt werden (vgl. z. B. Altepost 2021).

## 5. Literatur

- Altepost A (2021) Das kann man ja auch mal so sehen... Multiperspektivisches Assessment in der KI-Arbeitsplatzgestaltung. In: Gesellschaft für Arbeitswissenschaft e. V. (Hrsg.): Arbeit HumAIIn gestalten. Konzepte menschenzentrierter KI-Arbeitsplätze - Jetzt für die Arbeit von morgen vordenken. Bericht zum 67. Kongress der Gesellschaft für Arbeitswissenschaft 03. bis 05.03.2021, Bochum. – Dortmund: GfA-Press.
- Altepost A, Merx W, Rezaey A, Löhner M (2021) Neue Soziotechnische Systeme in der Textilbranche. Mensch-Technik-Interaktion und Ergebnisse der partizipativen soziotechnischen Systemgestaltung für ein Assistenzsystem in der deutschen Textilindustrie. Institut für Textiltechnik und Institut für Soziologie der RWTH Aachen University, Aachen.
- Häußling R, Ziesen N, Lemm J, Kerpen D, Strüver N, Eggert M (2017) Schlaglichter der Digitalisierung: Virtureale (r) Körper-Arbeit-Alltag: Ein Vorstoß zum Kern der Digitalisierung aus einer techniksoziologisch-relationalen Perspektive: Working Paper des Lehrstuhls für Technik-und Organisationssoziologie (No. RWTH-2017-06217). Lehrstuhl für Soziologie mit dem Schwerpunkt Technik-und Organisationssoziologie.
- Schmidt J, Hennemann J, Müller F, Thönis M, Bahlow J, Kötter W, Roth S, Lennings F, Terstegen S, Beck I, Schlink B, Vollborth T, Bendel A, Latniak E, Bartels E (2020) Arbeits- und prozessorientiert digitalisieren. Vorgehensweisen, Methoden und Erfahrungen aus dem Projekt APRODI. Eschborn: RKW.
- Stowasser S, Suchy O, Müller N (2020) Einführung von KI-Systemen in Unternehmen. Gestaltungsansätze für das Change-Management. Whitepaper aus der Plattform Lernende Systeme, München.
- Strohm O, Ulich E (Hrsg.) (1997) Unternehmen arbeitspsychologisch bewerten: ein Mehr-Ebenen-Ansatz unter besonderer Berücksichtigung von Mensch, Technik und Organisation (Vol. 10). vdf Hochschulverlag AG.
- VDI/VDE-Richtlinie 3714 (2020) Implementierung und Betrieb von Big-Data-Anwendungen in der produzierenden Industrie. Durchführung von Big-Data-Projekten. Berlin: Beuth Verlag.

**Danksagung:** Wir danken dem Bundesministerium für Bildung und Forschung für die Förderung und dem Projektträger Karlsruhe für die Betreuung des Kompetenzzentrums WIRKsam (Förderkennzeichen: 02L19C601) im Rahmen der Fördermaßnahme „Regionale Kompetenzzentren der Arbeitsforschung“. Ebenfalls danken wir unseren Forschungspartnern ifaa – Institut für angewandte Arbeitsforschung, Düsseldorf, und MASKOR Mobile Autonome Systeme und Kognitive Robotik, Fachhochschule Aachen, sowie unseren Unternehmenspartnern in WIRKsam.



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)