



Gesellschaft für
Arbeitswissenschaften



Institut für
Arbeitswissenschaft



Lehrstuhl für
Wirtschaftspsychologie



RUHR-UNIVERSITÄT BOCHUM

Uta Wilkens, Christian Cost Reyes, Tim Treude, Annette Kluge

Understandings and perspectives of human-centered AI. A transdisciplinary literature review

Problemstellung und Zielsetzung

„Menschenzentrierte KI“ / „human-centered AI“ häufig und multi-disziplinär verwendet,
aber vielfältige Begriffsverständnisse

Ziel:

Spektrum der Begriffsverwendung explizieren und
hinsichtlich der Implikation für die Arbeitsgestaltung einordnen

Methodik: Systematische Literaturanalyse

Suche nach keywords in den Literaturdatenbanken “Google Scholar” und “EBSCOhost” (Deutsch und Englisch)

“human-centered” oder “people-centered” und “artificial intelligence” oder “AI” oder “machine learning”

Publikationstyp: Zeitschriftenbeiträge, Bücher, Buchbeiträge, Forschungsberichte

Disziplinen: Arbeitswissenschaft, Ingenieurwissenschaften Information Science, Medizin, Psychologie, Management, Philosophie, Human factor studies, Soziologie, Erziehungswissenschaft etc.

⇒ 133 Referenzen identifiziert, davon 85 für die weitere Inhaltsanalyse

⇒ hermeneutische Auswertung: Clusterung nach der in der Quelle erkennbaren impliziten oder expliziten Grundannahme darüber, wann KI menschenzentriert ist; welches (implizite) Menschenbild kommt in den Ausführungen zum Vorschein?

(Überführung in weitere keywords wie “help”, “protection”, “reliability”, “potential”, “regulation” etc. kann erst im nächsten Analyseschritt erfolgen)

Perspektiven auf menschenzentrierte Arbeit mit und durch KI

Ergebnis der Inhaltsanalyse von 85 Quellen

Was ist es (derzeit) nicht, weil nicht auf menschenzentrierte Arbeitsgestaltung gerichtet?



Bildquelle: 3Dnatives

Humanoide Roboter

Ahrens 2014; Azvine & Wobcke 1998; Bettoni 1995



Bildquelle: imaginima

Mensch als Messmodell und Datenlieferant

Fiebrink & Gillies 2018; Gillies et al. 2016

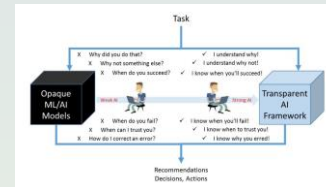


Bildquelle: Automationspraxis

1. Defizit-orientierte Perspektive

Ausgleich menschlicher Schwächen in Aufmerksamkeit, Konzentration, körperlicher Belastbarkeit, Ermüdungserscheinungen (moderner Taylorismus)

Adnan et al. 2020; Albu-Schäffer 2019; Bond et al. 2019; Djian et al. 2000; Oviatt 2006; Peruzzini & Pellicciari 2017; Romero et al. 2016; Schaal 2007; Schmidt 2020; Schmidtler et al. 2015; Tombeil et al. 2020



Bildquelle: sogetiLabs

2. Perspektive der Datenverlässlichkeit

Datenmuster sind unverzerrt und diskriminierungsfrei, Schlüsse sind nachvollziehbar und selbsterklärend, Daten sind vertrauenswürdig

Adadi & Berrada 2018; Barredo et al. 2020; Bond et al. 2019; Dewey & Wilkens 2019; Ehsan & Riedl 2020; Garcia-Magarino et al. 2019; Gunning et al. 2019; Hayes & Moniz 2020; How 2019; How und Chan 2020; How et al. 2020a; How et al. 2020b; Kahng 2019; Lukowicz 2019; Mucha et al. 2019; Ploug & Holm 2020; Ribera & Lapedriza 2019; Riedl 2019; Springer 2019; Steels 2020; Wang et al. 2019b; Xu 2019; Zanzotto 2019; Zhang et al. 2020; Zhu et al. 2018



Bildquelle: GesundheitsProfi

3. Perspektive der Arbeitsschutzes

Ergonomie und Job Design

Buxbaum & Häusler 2020; Elkmann 2013; Freitag et al. 2016; Häusler & Sträter 2020; Lam 2011; Matthias et al. 2013; Montreuil et al. 2007; Mutlu & Forlizzi 2008; Riemer et al. 2005; Riemer et al. 2006; Scholtz 2003; Steels 2020



Bildquelle: © Charité

4. Potenzial-orientierte Perspektive

Integration menschlicher und künstlicher Intelligenz, Komplementarität um menschliche Fähigkeiten weiter zu stärken, human-in-the-loop

Ahrens 2014; Akata et al. 2020; Bartlett 2006; Beede et al. 2020; Botthoff 2015; Böhmman et al. 2020; Cui & Dai 2008; Dewey & Wilkens 2019; Guszczka et al. 2017; Guszczka 2018; Holzinger et al. 2019; Huchler 2015; Huchler 2016a, 2016b; Huchler & Rhein 2017; Jarrahi 2018; Kaiser & Malanowski 2019; Nahavandi 2019; Nichols et al. 2015; Nisser & Malanowski 2019; Riedl 2019; Roesler & Onnasch 2020; Romero et al. 2016; Schirner et al. 2013; Schmidt 2020; Shneiderman 2020a, 2020b, 2020c; Wang et al. 2019a; Wilkens et al. 2019; Wilkens 2020; Wilson & Daugherty 2018



gg101940270 www.gopgraph.com

5. Perspektive der Machtverteilung zwischen Mensch und KI

Informationszugänge und Informationsverteilung, neue Organisationsmodelle bis hin zur Regulation, Unterordnung der KI unter Menschen als normatives Kriterium

Carrico, 2018; Cimini et al., 2020; Hinds et al. 2004; Kaiser und Malanowski 2019; Pacaux-Lemoine et al. 2017; Schaal 2007

Diskussion: Perspektiven für die Arbeitsgestaltung

**Hinreichende
Bedingungen:**



**1. Defizit-orientierte
Perspektive**



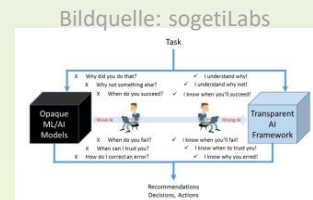
**3. Perspektive der
Arbeitsschutzes**



**4. Potenzial-orientierte
Perspektive**

**Notwendige
Bedingungen:**

**2. Perspektive der
Datenverlässlichkeit**



KI-intern

**5. Perspektive der Machtverteilung
zwischen Mensch und KI**



KI-extern



Gesellschaft für
Arbeitswissenschaften



Institut für
Arbeitswissenschaft



Lehrstuhl für
Wirtschaftspsychologie



RUHR-UNIVERSITÄT BOCHUM

Einladung zur Diskussion: Kontaktdaten

Ruhr-Universität Bochum
Institut für Arbeitswissenschaft
Lehrstuhl Arbeit, Personal und Führung
44780 Bochum

www.apf.rub.de

Prof. Dr. Uta Wilkens
uta.wilkens@rub.de

Christian Cost Reyes
christian.costreyes@rub.de

Tim Treude
tim.treude@rub.de

Ruhr-Universität Bochum
Fakultät für Psychologie
Lehrstuhl für Wirtschaftspsychologie
44780 Bochum

<https://www.aow.ruhr-uni-bochum.de>

Prof. Dr. Annette Kluge
annette.kluge@rub.de