

Widersprüchliche Arbeitsanforderungen bei Koaktion von Fachleuten und adaptiven Systemen

Peter Brödner

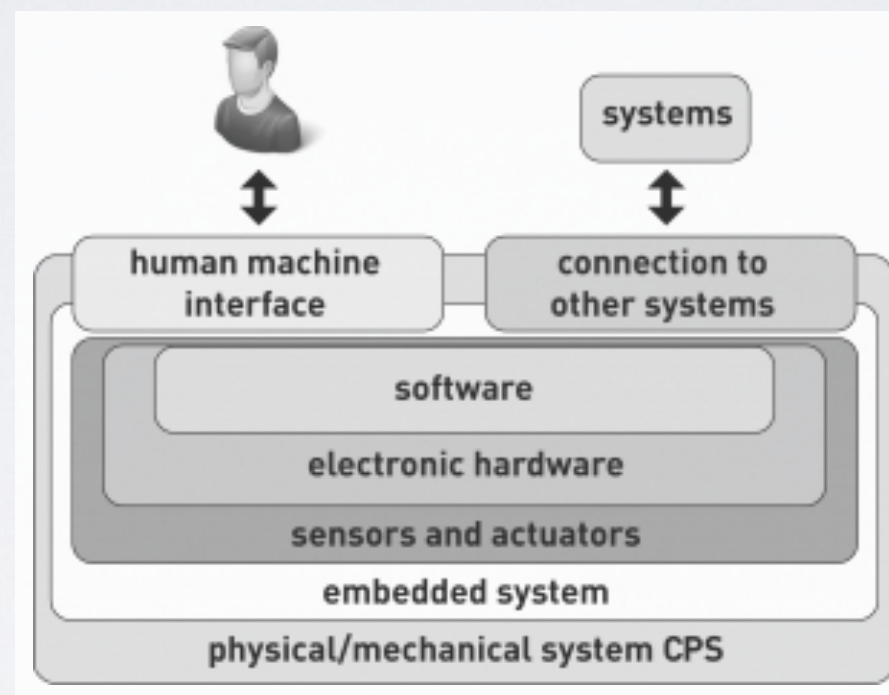
Tagung der Leibniz-Sozietät der Wissenschaften zu Berlin
**Zukunft der Arbeit – soziotechnische Gestaltung der Arbeitswelt
im Zeichen von »Digitalisierung« und »Künstlicher Intelligenz«**

Berlin, 13. Dezember 2019

Kennzeichen »intelligenter« Agentensysteme

Anders als ihr Name suggeriert, sind »intelligente Agenten« als fortgeschrittene Computersysteme lediglich **kontextsensitiv** oder **adaptive** Systeme, die automatisch auf bestimmte Umweltbedingungen rational zu reagieren vermögen (Russell & Norvig 2009).

Dazu verfügen sie über eine **Mensch-Maschine-Schnittstelle**, können über **Sensoren** Signale aus ihrer Umgebung aufnehmen, über **Aktoren** auf diese einwirken und **Daten** mit anderen Systemen **austauschen**. Sie arbeiten programmgesteuert **auto-operational** mittels berechenbarer Funktionen.

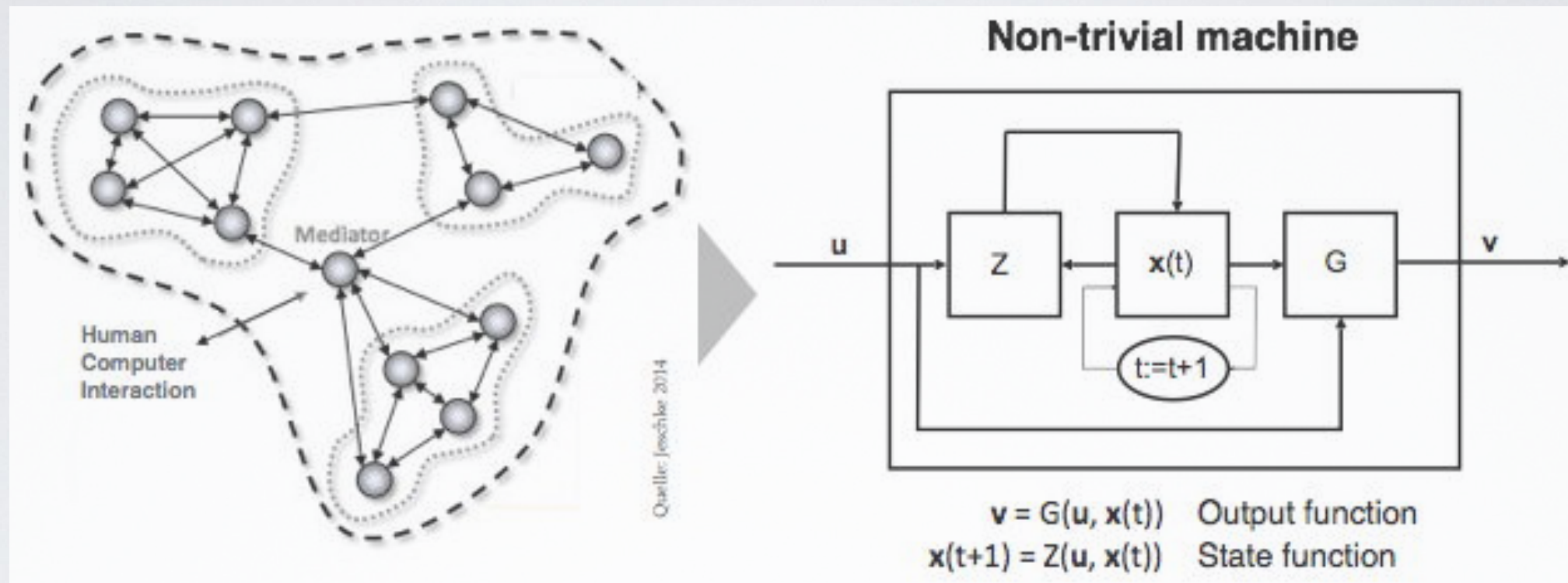


(Source: According to Broy 2010)

Irrtümlich werden diese Systeme häufig auch als »**autonom**« oder »**selbstlernend**« bezeichnet. Tatsächlich sind sie nicht autonom, da ihre Operationsregeln **nicht selbstbestimmt**, sondern von Ihren **Entwicklern eingeschrieben** sind. Als per Algorithmen programmgesteuerte Systeme sind sie nicht »intelligenter« als jedes andere Computersystem, dazu befähigt »Aufgaben zu bewältigen, die nach landläufiger Auffassung Intelligenz erfordern« (Autorengruppe 2018).

Multi-Agenten-Systeme (MAS) als »nicht-triviale Maschinen«

Adaptive Systeme wie auto-operationale »Agenten« können zu **Multi-Agenten Systemen (MAS)** verknüpft werden. Das **Verhalten von MAS** hängt stark von der **Wechselwirkung** mit ihrer sich ändernden Umgebung, ihr jeweiliger Zustand damit von der **Vorgeschichte** ab. Formal gleicht es dem Verhalten »**nicht-trivialer Maschinen**« (von Foerster 1993).



Aufgrund der Wechselwirkung mit der Umgebung und der Geschichtsabhängigkeit ist das Verhalten von MAS, obgleich durch Algorithmen **vollständig determiniert**, tatsächlich mehr oder weniger **undurchschaubar** und **nicht nachvollziehbar**.

Das Verhalten ist, mit anderen Worten, **situiert**, **geschichtsabhängig** und für Nutzer **nicht vorhersehbar**.

Arbeit mit adaptiven Systemen: **Koaktion statt Interaktion**

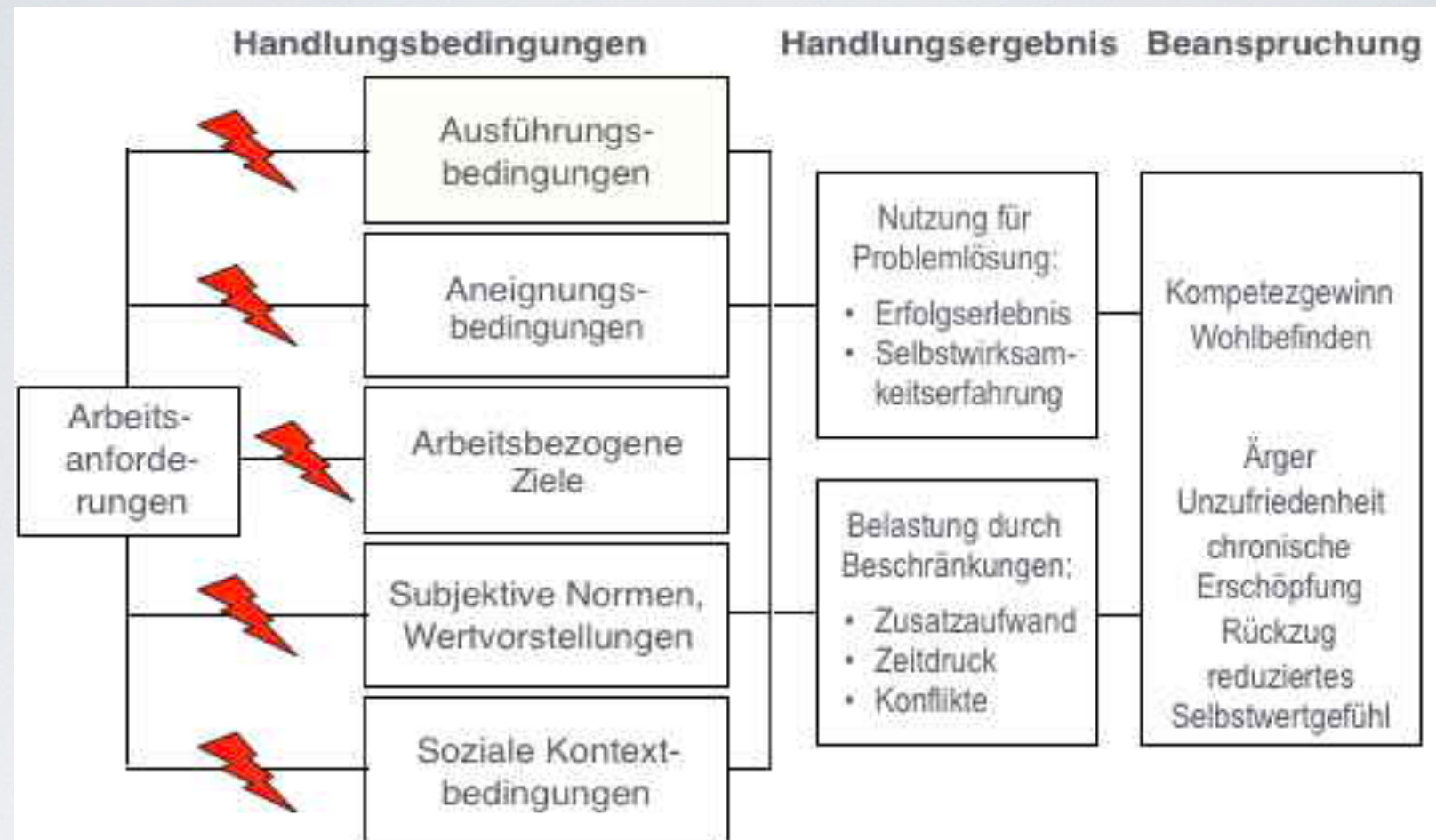
Auch wenn sich **adaptive Systeme** wie MAS, die sich **algorithmisch gesteuert** veränderlichen **Umweltbedingungen anpassen**, verbreiten, ist auf absehbare Zeit eher **nicht zu erwarten**, dass diese – entgegen oftmals illusionären Erwartungen – qualifizierte (Wissens-)Arbeit vollständig automatisieren und **ersetzen**. Folglich sind qualifizierte (Wissens-)Arbeiter dazu gezwungen, bei der Bewältigung ihrer Aufgaben mit derartigen Systemen zu »**koagieren**« (Hubig 2019).

Koaktion intentional handelnder Menschen mit **adaptiven Systemen** tritt dann an die Stelle instrumenteller Mensch-Maschine-**Interaktion** durch hinreichende Aneignung der Systemfunktionen. Dagegen ist intentionale, **zielgerichtete Interaktion** mit adaptiven Systemen, ihr **instrumenteller Gebrauch**, wegen des **situierten** und **undurchschautes Verhaltens** kaum möglich (vgl. Bradshaw et al.2011).

Dadurch sehen sich qualifizierte (Wissens-)Arbeiter im Umgang mit **adaptiven Systemen** vor neue **Herausforderungen** und **Belastungen** gestellt:

- Die effektive Bewältigung komplexer Arbeitsaufgaben unterliegt – bei hohen Erfolgserwartungen – dauerhaft möglichen Störungen durch ungewöhnliches, undurchschautes Systemverhalten.
- Es ist kaum noch möglich, die situierten Systemreaktionen nachzuvollziehen oder zu erklären.
- Dadurch werden (Wissens-)Arbeiter daran gehindert, aus Erfahrung zu lernen und sich wichtige Systemfunktionen hinreichend anzueignen.
- Trotz dieses Kontrollverlusts werden sie oft für eingetretene Fehler und Schäden verantwortlich gemacht.
- Durch derartige **widersprüchlichen Arbeitsanforderungen** werden (Wissens-)Arbeiter insgesamt erheblichen **Belastungen** ausgesetzt, die auf Dauer **psychische Störungen** hervorrufen können.

Psychische Störungen durch widersprüchliche Arbeitsanforderungen



Quellen: Moldaschl 2005; Brödner 2009; Gerlmaier & Latniak 2011

Eine relational ressourcenbasierte Sicht auf Arbeit und operationales Modell der Stressgenese

Handlungsbedingungen können je nach situativem Kontext als **Beschränkung** oder **Ressource** wirken.

Beispiele für belastende Handlungsergebnisse:

- Widersprüche zwischen **Aufgaben** und **Ausführungsbedingungen** oder **Lernmöglichkeiten** beschränken die Handlungsregulation bzw. die Aneignung nötigen Wissens: z.B. Zusatzaufwand durch unpassende Werkzeuge,
- Widersprüchliche **Projektziele** stürzen Arbeitspersonen in Loyalitätskonflikte,
- Widersprüche zwischen **aufgabenbezogenen** & **persönlichen Werten** verursachen Wertkonflikte,
- Widersprüche zwischen **erwartetem** & **tatsächlichem** Verhalten »intelligenter **Agenten**« verursachen unter hoher Leistungserwartung beträchtlichen Zusatzaufwand sowie Loyalitäts- & Wertkonflikte.

Wissenschaftliche und ethische Herausforderungen

Im Umgang mit **adaptiven Systemen** sind die nachstehenden **Herausforderungen** zu **bewältigen**. Ihre **Lösung** erscheint umso **dringlicher**, je komplexer und damit auch undurchschaubarer die Systeme sind und je anfälliger für Störungen oder Missbrauch sie werden (Mittelstadt et al. 2016):

- »**Inconclusive evidence**«: Das Systemverhalten ist **intransparent**, selbst für Entwickler.
- »**Inscrutable evidence**«: Das Systemverhalten ist kaum **erklärbar**, selbst im Rückblick.
- »**Misguided evidence**«: Systeme unterliegen unbekanntem **Störungen** und erzeugen **nur wahrscheinliche**, keine sicheren **Ergebnisse**, die zudem auf **Daten** zweifelhafter, **nicht einschätzbarer Qualität** beruhen.
- »**Transformative effects**«: Nutzer werden zu übertriebenen, **unrealistischen Erwartungen** über Systemleistungen verleitet.

Nutzer sind infolge der Intransparenz der Systeme

- zu **blindem Vertrauen** in die Systemausgaben verdammt, deren Qualität sie nicht beurteilen können,
- daran **gehindert**, sich die Systemfunktionalität für instrumentelles Handeln hinreichend **anzueignen** (Verletzung des Grundsatzes der »Ewartungskonformität« (EN ISO 9241-11)),
- im **Unklaren** darüber gelassen, wer für **Fehlleistungen verantwortlich** gemacht wird.

Nach dem **Modell** der **Stressgenese** werden (Wissens)Arbeiter – konfrontiert mit derart widersprüchlichen Arbeitsanforderungen, unter **Leistungsdruck** und **ohne volle Kontrolle** über ihr **Arbeitsmittel** – hohen Risiken psychischer Störungen ausgesetzt (vgl. auch Gerlmaier & Latniak 2011, Norman 1994).

Darüber hinaus werden wohl bekannte »**Ironien der Automatisierung**« auf die Spitze getrieben: Trotz im Normalbetrieb unzureichend entwickelter Kompetenz und Erfahrungen würden diese im Störfall in hohem Maße benötigt; die Folge können schwere Fehlleistungen sein (Bainbridge 1983; Baxter et al. 2004).

Empfehlungen für die soziotechnische Arbeitsgestaltung

Gestützt auf das vorgestellte **operationale Modell der Stressgenese** können **Schlussfolgerungen** zur **soziotechnischen Arbeitsgestaltung** gezogen werden. Um **produktive** und **zugleich sozialverträgliche** Arbeitssysteme zu schaffen, in denen (Wissens-)Arbeiter durch fortgeschrittene Computersysteme **unterstützt** werden, sind folgende **Anforderungen** zu **erfüllen**:

- Um **widersprüchliche Arbeitsanforderungen** zu vermeiden, sollte der Einsatz adaptiver Systeme auf solche Aufgaben beschränkt werden, die sich erfolgreich **vollständig automatisieren** lassen.
- Für Aufgaben, deren Bewältigung weiterhin qualifizierte (Wissens-)Arbeit erfordert, sollte der **Einsatz adaptiver Systeme** aus den dargelegten Gründen **vermieden** werden (solange die Systeme ihr Verhalten nicht selbst erklären können).
- **Forschungsanstrengungen** sollten sich auf Entwicklung fortgeschrittenener Computersysteme ausrichten, die qualifizierte (Wissens-)Arbeit **unterstützen** und **effektiver machen**, nicht ersetzen (vgl. auch Norman 2017; Brödner 2018).
- **Forschungsanstrengungen** zu adaptiven Systemen sollten ferner auf die Entwicklung und Implementierung von Einrichtungen konzentriert werden, die auf Verlangen fragwürdiges **Systemverhalten** nachvollziehbar zu **erklären** vermögen (oft gefordert, bei weitem nicht erreicht).
- Wie andere **Hochrisiko-Artefakte** sollten auch **adaptive Systeme öffentlich kontrollierter Zertifizierung** unterliegen.
- In Fällen des **Systemversagens** sollte Verantwortung nur aufgrund **sorgfältiger Untersuchungen** und **umfassender Analysen**, nicht vorschnell menschlichem Versagen, zugeschrieben werden. Dafür sind ausreichende Untersuchungskapazitäten einzurichten.

Literatur

Autorengruppe (2018): The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation, Oxford (AR): Future of Humanity Institute u.a. 02/2018, <https://arxiv.org/pdf/1802.07228.pdf>

Bainbridge, L. (1983): Ironies of Automation, *Automatica* 19 (6), 775-779

Baxter, G.; Rooksby, J.; Wang, Y. & Khajeh-Hosseini, A. (2012): The Ironies of Automation ... still going strong at 30?, in: P. Turner & S. Turner (eds.): European Conference on Cognitive Ergonomics, ECCE '12, Edinburgh (UK), August 28-31, 2012, 65-71

Bradshaw, J.M.; Feltovich, P. & Johnson, M. (2011): Human-Agent Interaction, in: Boy, G.A. (ed.): The Handbook of Human-Machine Interaction. A Human-Centered Design Approach, Boca Raton (FL): CRC Press

Brödner, P. (2018): »Super-intelligent« Machine: Technological Exuberance or the Road to Subjection, *AI & Society Journal of Knowledge, Culture and Communication* 33, pp. 335–346

Brödner, P. (2009): Sustainability in Knowledge-Based Companies, in: Docherty, P.; Kira, M. & Shani, R. (eds.) (2009): Creating Sustainable Work Systems. Developing Social Sustainability, London: Routledge, 53-69

Foerster, H. von (1993): Wissen und Gewissen, Frankfurt/M: Suhrkamp

Gerlmaier, A. & Latniak, E. (Hg.) (2011): Burnout in der IT-Branche. Ursachen und betriebliche Prävention, Kröningen: Asanger Verlag

Hubig, C. (2019): Haben autonome Maschinen Verantwortung?, in: Hirsch-Kreinsen, H. & Karacic, A. (Hg.): Autonome Systeme und Arbeit. Perspektiven, Herausforderungen und Grenzen der Künstlichen Intelligenz in der Arbeitswelt, 275-298

Mittelstadt, B.D.; Allo, P.; Taddeo, M.; Wachter, S. & Floridi, L. (2016): The Ethics of Algorithms: Mapping the Debate, *Big Data & Society* 3 (2), 1-21

Moldaschl, M. (ed.) (2005) Immaterielle Ressourcen. Nachhaltigkeit von Unternehmensführung und Arbeit I, München: Hampp

Norman, D (2017): Design, Business Models, and Human-technology Teamwork, *Research-Technology Management*, 60 (1), 26-29

Norman, D.A. (1994): How Might People Interact with Agents, *CACM* 37 (7), 68-71

Russell, S. & Norvig, P. (2009): Artificial Intelligence: A Modern Approach, 3rd. ed., Upper Saddle River (NJ): Pearson