

SocietyByte

BFH-Magazin für die Humane Digitale Transformation

Drei Facetten der humanen digitalen Transformation

Von Mascha Kurpicz-Briki (BFH Technik & Informatik) | 0 Kommentare



Wie wird die digitale Transformation unsere Gesellschaft und unser Leben beeinflussen, und was müssen wir dabei beachten, so dass der Mensch im Zentrum bleibt? Unsere Autorin geht auf drei Facetten ein.

Die digitale Transformation schreitet schnell voran, und hat das Potential, die Art und Weise wie wir leben, arbeiten und lernen zu verändern. Insbesondere mit Technologien aus dem Bereich des maschinellen Lernens kommt es zu neuen Herausforderungen. Was gilt es zu beachten, damit dieser Änderungsprozess menschenfreundlich bleibt?

Das neue strategische Themenfeld humane digitale Transformation [<https://www.bfh.ch/de/strategische-themenfelder/digitalisierung/>] der Berner Fachhochschule setzt sich genau mit diesem Thema auseinander. Das Themenfeld beschäftigt sich damit, wie die Technologie am besten und nachhaltig genützt werden kann, und der Mensch und seine Bedürfnisse ins Zentrum gestellt wird. Doch was bedeutet dies für die Entwicklung von Technologie?

Dieser Artikel soll drei solche Aspekte beleuchten, welche sich an den folgenden Leitfragen orientieren:

1. Wie muss die Zusammenarbeit zwischen Menschen und Technologie gestaltet

- sein, so dass ein verantwortlicher und akzeptabler Einsatz möglich ist?
2. Welche ethischen und menschen-bezogenen Faktoren sind relevant für das Training von KI-Modellen?
 3. Was ist zu beachten bezüglich Stereotypen in den Trainingsdaten und potenzieller Diskriminierung durch KI-Entscheidungen?

Wir und die Technologie

In der aktuellen Diskussion wird die künstliche Intelligenz leider regelmässig als allwissender Roboter dargestellt, der es sich irgendwann zum Ziel nimmt, die Menschheit zu zerstören. Dies passiert sowohl in den breiten Medien, als auch in einem Teilgebiet der Fachdiskussion. Die Meinungen, ob es je zu so einem Szenario kommt, und falls ja, wann, gehen auseinander. Dieses Ziel, eine *general artificial intelligence* zu entwickeln, welche mehr einem Hollywood Skript entspricht als einem nützlichen Tool, muss jedoch in Hinblick auf eine humane digitale Transformation hinterfragt werden.

Es gibt viel dringendere Problemstellungen in der Zusammenarbeit zwischen Menschen und Technologie, welche bereits bei heute existierenden Tools relevant sind. Einige Forscherinnen der BFH riefen in einem Beitrag [1] dazu auf, insbesondere die Herausforderungen, wie Menschen und Maschinen zusammenarbeiten, anzugehen. Ein wichtiges Beispiel ist hier die Tatsache, ob ein Mensch von der Maschine (oder der Software) Befehle erhält, oder der Mensch diese Technologien als Werkzeuge in der Arbeitsunterstützung nutzt, aber die Kontrolle über das Geschehen und die Reflektion der Entscheidungen beim Menschen bleibt.

Wir setzen uns daher dafür ein, dass solche Technologien im Rahmen einer *erweiterten* statt *künstlichen* Intelligenz umgesetzt werden (augmented statt artificial intelligence). Statt den Menschen komplett zu ersetzen, sollen die Technologien als schlaue Werkzeuge den (Arbeits-)Alltag des Menschen unterstützen. Dies ist einerseits nötig, weil nur der Mensch den Gesamtkontext im Blick behalten kann, und entsprechend Entscheidungen oder Vorschläge kritisch reflektieren. Die KI ist nur so gut wie ihre Trainingsdaten und arbeitet mit Wahrscheinlichkeiten. Dies mag oftmals gut funktionieren, ist jedoch in kritischen Entscheidungen über Menschen oder Ressourcen nicht ausreichend.

Die Menschen hinter den Kulissen

Ein weiterer, oft unterschätzter Punkt der Menschlichkeit in Bezug auf KI-Systeme sind die Menschen, welche in den Trainingsprozessen involviert sind. Damit solche Systeme lernen können, werden oft Menschen eingesetzt, beispielsweise um Texte von Hand in Kategorien einzuordnen. Während dies teilweise unproblematisch sein kann, besteht in bestimmten Anwendungsfällen und bei schlechten Arbeitsbedingungen grosses Diskussionspotential.

Eine Recherche des TIME Magazins [2] deckte genau dies in Zusammenhang mit ChatGPT auf. Um toxische und beleidigende Texte in den Antworten des bekannten Chatbots zu vermeiden, hatte das Unternehmen OpenAI diese Aufgabe an ein Unternehmen in Kenya ausgelagert. Dort mussten teils verstörende Texte für rund 2 Dollar die Stunde durch Menschen bearbeitet werden.

Eine weitere Frage in diesem Zusammenhang ist die Herkunft der Daten. Wer hat die Texte geschrieben, aus denen mittels KI neue Werke geschaffen werden? Ähnlich verhält es sich mit Bildern und Sprache, bei denen die Trainingsdaten der KI von menschlichen Künstler*innen geschaffen wurde. Der Umgang mit dieser Problematik bietet viele offene Fragen für die digitale Gesellschaft der Zukunft.

Fairness und Bias

Der dritte Aspekt der menschlichen Komponenten der digitalen Transformation umfasst das Problem der Diskriminierung durch KI. Viele Beispiele haben in den vergangenen Jahren gezeigt, wie solche Technologien die Stereotypen der Gesellschaft abbilden. Dies hat einen Einfluss auf die Ergebnisse solcher Modelle. In einem bekannten Beispiel wurde eine Software trainiert, um Bewerbungsdossiers zu sortieren [3]. In einem vorherigen SocietyByte Artikel hatten wir über Stereotypen in ChatGPT berichtet [4]. Dieses Problem technisch zu lösen ist sehr herausfordernd und noch Gegenstand der laufenden Forschung. Insbesondere ist die Entwicklung solcher Erkennungsmethoden für die verschiedenen Sprachen herausfordernd, und es kann kulturelle Unterschiede geben, bezüglich der Stereotypen, welche vorhanden sind [5][6].

Über die AI-Forschung an der BFH

Die Forschungsgruppe Applied Machine Intelligence [<https://www.bfh.ch/en/research/research-areas/applied-machine-intelligence/>] setzt sich mit den wissenschaftlichen, technischen und gesellschaftlichen Herausforderungen von KI-Technologien auseinander. Mit einem angewandten Fokus setzen sie den Menschen ins Zentrum. Sie sind unter anderem der technische Partner im EU-Projekt BIAS [<http://www.biasproject.eu>], welches erforscht, wie die Diskriminierung von KI-Anwendungen im Bereich HR erkannt und reduziert werden kann.

References

- [1] <https://www.frontiersin.org/articles/10.3389/frobt.2022.997386/full> [<https://www.frontiersin.org/articles/10.3389/frobt.2022.997386/full>]
- [2] <https://time.com/6247678/openai-chatgpt-kenya-workers/> [<https://time.com/6247678/openai-chatgpt-kenya-workers/>]
- [3] <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G> [<https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>]
- [4] <https://www.societybyte.swiss/2022/12/22/hi-chatgpt-hast-du-vorurteile/> [<https://www.societybyte.swiss/2022/12/22/hi-chatgpt-hast-du-vorurteile/>]
- [5] <https://www.frontiersin.org/articles/10.3389/fdata.2021.625290/full> [<https://www.frontiersin.org/articles/10.3389/fdata.2021.625290/full>]
- [6] <https://ceur-ws.org/Vol-2624/paper6.pdf> [<https://ceur-ws.org/Vol-2624/paper6.pdf>]



AUTHOR: MASCHA KURPICZ-BRIKI



Dr. Mascha Kurpicz-Briki ist Professorin für Data Engineering am Institute for Data Applications and Security IDAS der Berner Fachhochschule, und stellvertretende Leiterin der Forschungsgruppe Applied Machine Intelligence. Sie beschäftigt sich in ihrer Forschung unter anderem mit dem Thema Fairness und der Digitalisierung von sozialen und gesellschaftlichen Herausforderungen.

Posts from Mascha Kurpicz-Briki

Create PDF

Ähnliche Beiträge

Humane digitale Transformation: Auf dem Weg zu einer menschenzentrierten Digitalisierung

0

COMMENTS