Critical Reflection of AI Applications for persons with Disabilities in Vocational Rehabilitation

Susan Beudt[1], Berit Blanc[2], Rolf Feichtenbeiner[3] and Marco Kähler[4]

Abstract: Applications of artificial intelligence are increasingly being used to support work and learning in the workplace. Adaptivity and recommender systems, as key features of such innovative technologies, allow for enhanced personalization. Most notably, persons with disabilities may benefit from such technologies at work and during on-the-job training. Adapting such systems to very heterogeneous target groups, however, is not easily done. Implementing AI-based assistive systems in various educational settings in vocational training, especially in vocational rehabilitation, can also be challenging. This position paper looks at existing AI-based applications to analyze their potential for more inclusive workplaces and qualification processes. Furthermore, those technologies are discussed in the context of current ethical discourses to identify to what extent normative requirements are being reflected in existing AI-based applications.

Keywords: assistive technologies, inclusion, AI, ethics, vocational training

1 Introduction

More and more digital tech is used at the workplace and to support vocational training, e.g., in the form of assistance and knowledge services on the shop floor and the office floor. Those systems, however, are rarely designed and developed for a heterogeneous and diverse user group. Awareness of this issue is slowly but increasingly changing. Assistive Technology has been used to support persons with disabilities for a long time. The focus here is often put on physical disabilities enabling assistive technology - although often not AI-based - to remove at least some barriers at the workplace or during vocational training.

AI-based assistive technology may be one key component when aiming at removing barriers for a diverse target group. Through their ability to adapt automatically to the needs of the individual user, such systems allow for highly personalized assistance. We, therefore, see great potential to enhance the participation of persons with disabilities at the workplace. Those technologies may enable persons with disabilities to use their strengths and skills without being held back by various barriers in individual work environments.

[1] Deutsches Forschungszentrum für Künstliche Intelligenz, Educational Technology Lab, Berlin, susan.beudt@dfki.de
[2] Deutsches Forschungszentrum für Künstliche Intelligenz, Educational Technology Lab, Berlin, berit.blanc@dfki.de
[3] Deutsches Forschungszentrum für Künstliche Intelligenz, Educational Technology Lab, Berlin, rolf.feichtenbeiner@dfki.de
[4] Deutsches Forschungszentrum für Künstliche Intelligenz, Educational Technology Lab, Berlin, marco.kaehler@dfki.de

doi:10.18420/delfi2020-ws-116

In vocational rehabilitation, where diverse learning and working settings exist, those technologies may open up new employment opportunities or support reintegration in working life (e.g., enabling new work areas).

While opportunities seem promising, one must carefully analyze and evaluate potential risks that can emerge through implementing and using respective AI applications for persons with disabilities for vocational training and vocational rehabilitation. For the last years, the overall debates about AI have shifted towards ethical issues. The debates were essential to lay out the key aspects of AI ethics. So far, however, ethical discussions revolved mainly around a relatively homogeneous target group, in which persons with disabilities were only rudimentarily brought into focus.

With this positional paper, we want to take a closer look at existing AI-based assistance systems that are either developed directly for persons with disabilities or have the potential to support them at work. Thereby, we will focus on factors that are currently underrepresented in the AI ethics debate. This input may guide the design and development of future AI-based assistance systems. We also looked at the information given by technology providers on data sovereignty and data protection. Many evolving initiatives discuss the need for standardization of AI applications [Zi20]. Based on first results of our research project KI.ASSIST (on AI-based assistance systems for persons with severe disabilities in vocational rehabilitation), we point out the importance of transparency and standardization in the presentation of such systems by the manufacturers or technology providers.

2 AI Ethics and inclusion

The last years have seen a growing interest in ethics revolving around AI-based technologies. To date, a considerable number of ethical guidelines have been put forward by key political and economic organizations, the private sector, stakeholders from science, and the non-profit sector. Most frameworks discuss standards concerning the design, development, and application contexts. Main ethics guidelines have been introduced by the European Commission's High-Level Expert Group on AI [AI19], the OECD [OE20], the German Data ethics commission [DE19] or are currently developed such as, for instance, by the Study Commission "Artificial Intelligence – Social Responsibility and Economic, Social and Ecological Potential" [En18]. Some companies set themselves ethical frameworks (for an overview of current AI ethics frameworks by various stakeholders, see [Al19]). Comparing existing guidelines showed that several core topics or codes of practice are deemed relevant across frameworks, both in the public and private sector. [Ha20] and [Yu19] listed the most and repeatedly discussed factors concerning AI ethics. Among the top 6 main issues covered by 22 selected guidelines were privacy, fairness, accountability, transparency, safety, and the common good (grouped with sustainability and well-being) [Ha20].

In most frameworks, diverse and disabled user groups are significantly being underrepresented. As [Ha19] states, in areas such as diversity in AI, fostering of solidarity, respect for human autonomy, and AI for the common good, there is still a long way to go to achieve the ethical goals set. Our analysis on the significance of inclusion related topics regarding seven prominent ethical guidelines on AI also indicates that neither the disability dimension nor the social or cultural dimensions of inclusion have been sufficiently addressed (own work, not yet published).

One of the challenges in formulating ethical guidelines for AI applications is the possibility of unintended applications of AI systems (e.g., dual-use). Usually, the systems are being designed for a specific use case in a particular domain and with a predefined user group in mind. However, most AI systems are quite generic and may be transferable or customizable and, therefore, can be used in various contexts and for various purposes [Br18]. While inclusion by design is still an emerging field, persons with disabilities often have no choice but to use AI systems that are developed for a non-diverse user group. Thus, one critical point to make is that such systems could potentially cause harm beyond accessibility issues even when used for the purpose for which they were originally designed.

In our research, we looked at the extent to which ethical factors that are specifically relevant for inclusion and persons with disabilities are being integrated into the description of AI-based assistive technologies. Secondly, we also analyzed how current ethical aspects are being reflected in expert evaluations of such technologies.

3 Related Work: Scanning and Monitoring of AI-based Assistive Technologies

In many areas of the education and labor market, AI-driven digital transformation significantly shapes new forms of teaching, learning, and working. To assess the potential of AI applications to support vocational education and work practices for persons with disabilities, we set up a comprehensive and systematic technology foresight process. This process includes a scanning, monitoring, and scouting of AI-based assistive technologies to analyze and evaluate them. The aim is to build up a scientifically sound knowledge base that offers an orientation as to which AI-based assistance systems may add value for work and education scenarios in the context of occupational rehabilitation. The findings will also guide the testing of AI-based technologies in rehabilitation institutions and companies. In this way, it can also be investigated which groups could sustainably benefit from which AI applications. The findings will help us identify short-, medium- and long-term application possibilities and critical success factors for their implementation and formulate recommendations for the digital transformation of occupational rehabilitation systems. We present first results from our technology scanning and expert interviews focusing on aspects of ethics and data sovereignty.

3.1 Method

The methodological approach of the monitoring comprises three parts: research, validation, and foresighting. As a first step, we conducted systematic research of AI-based assistance systems for persons with disabilities in vocational rehabilitation. We scanned AI applications that represent finished products available on the market and technologies currently being developed for commercial and research purposes.

The basis for this comprehensive web research is a search, description, and evaluation system developed in the project network. Therefore, we defined search objects and search terms (English and German), search locations, selection and exclusion criteria, and description and evaluation categories to conduct a targeted search for AI-based assistive technologies that support work and training for persons with disabilities. The technology may have been developed in an inclusion context or should show potential for the support for persons with disabilities at the workplace. A comprehensive categorization system was created to provide the basis on which the suitability of technologies for persons with disabilities in vocational rehabilitation could be evaluated. This system initially contains numerous descriptive categories that were collected in the scanning of AI-based applications.

We conducted a summarizing and structuring qualitative content analysis of the technology features recorded in the desk research (following [Ma00], with a deviation regarding the source material). As part of the qualitative content analysis, assessments such as the development status of the application or ratings covering various forms of disability were derived from the description categories. Further evaluation categories were created to assess the benefit and suitability of the application for the target group. These evaluation categories were based on the seven person-centered dimensions formed in the project: competencies, acceptance, motivation, self-determination, data sovereignty, diversity, and participation. In a workshop with a multidisciplinary group of participants, these seven dimensions were subdivided into categories, from which evaluations were compiled in the content analysis, e.g., for personal added value or the acquisition effort.

In order to validate the suitability of the collected technologies for persons with disabilities at work, our analysis furthermore comprises three components: a qualitative content analysis of the scanned technologies, qualitative content analysis of guided interviews with inclusion as well as AI experts (according to [Ma00]) and a standardized online survey. In the final step of our monitoring activities, we will use a foresighting methodology to assess which trends and future developments are expected or probable and which are desired.

This paper presents work in progress and first results from desk research, qualitative content analysis, and expert interviews with a specific focus on aspects relating to ethics, data protection, and data sovereignty.

3.2 Results from desk research and qualitative content analysis

We found 154 AI-based assistive technologies that met the criteria described above. Fifty-nine of which were existing products, 89 non-finalized projects, and six demonstrators or prototypes. In two out of three technologies, types of disabilities are addressed, mainly physical (46) and sensory disabilities (38). Most of the collected technologies support visually impaired or blind people as well as people with hearing impairments. Learning disabilities (11) and mental disorders (9), on the other hand, are much less frequently addressed by the technologies.

The technologies were then systematized according to which tasks the technologies mainly perform. The applications were assigned to seven task groups: Perception and communication support, interaction support, navigation/orientation support, concrete work support, learning support, physical support, and psychological support. We also analyzed aspects related to ethics, data protection, and data sovereignty. An index to determine the suitability of these technologies for our target group and context was developed. This index was used to determine the most suitable technologies for each task group and randomly choose technologies for further analysis of ethics and data sovereignty. For further analysis, we selected four categories based on inclusion-relevant ethical aspects: self-determination, diversity, participation, and transparency. General AI-related ethical aspects such as accountability, technical robustness, and security have not been included.

Out of the 44 analyzed technologies, we found that the promotion of self-determination or independence from other people is part of most applications. Technology providers stated that with the help of their technology, people with disabilities could perform tasks entirely independently without needing help from other people (17) or perform subtasks more independently but still depend on other people's support (22).

With regard to diversity, we found that the majority of the providers stated that people with disabilities were actively considered in the training data of the AI system and that the technology was designed accessible (in 17 cases, both statements applied, in 12 cases, either one of the statements applied). In terms of participation, providers stated that persons with disabilities (in 13 cases) or their representatives (in four cases) were actively involved (at least consulted) in developing the technology.

We also looked at the extent to which the information made available by the providers of AI-based assistive technologies increases transparency for persons with disabilities, e.g., concerning the system, its functions, modes of operation, and possible malfunctions of the technology. For four out of 44 applications, such information was given in simple language, sign language (e.g., in videos), audio description, and Braille. In seven cases, it was only provided in one form.

For further analysis of the information given by technology providers concerning data protection and data sovereignty, we selected the following seven categories based on relevant sources (e.g. [De18], [Sa17]).

Out of the 62 analyzed technologies, we found the following results:

- **Data sovereignty - scope for action and decision-making** concerning data collection: For none of the technologies, information was provided on whether the user could choose between different data collection models/how much data is collected or whether sub-functions could be activated or deactivated.
- **Data sovereignty and self-determination**: In 10 cases, the user is informed about the personal data collected and can view and control (for example, view, delete or modify) the data. In three cases, the user is informed about the personal data collected or can view and control them. In the other 49 cases, the user is not informed about the collected personal data.
- **Data protection compliance**: In four cases, it is declared that the technology is data protection compliant. In one case, data protection conformity is clearly explained.
- **Data acquisition – description of personal data usage**: In nine cases, the collected data is named completely. In three cases, some areas/parameters are named, and for the other 50 technologies, no such information is mentioned.
- **Data processing**: In five cases, it is stated that collected personal data is only used anonymously and then deleted. In six cases, it is stated that the data is treated anonymously and pseudonymously, and for the other 50 technologies, no such information was found.
- **Data storage**: In four cases, it is stated that the collected data is only stored locally (storage exclusively within the EU is not indicated), while no information on the storage location can be found for the remaining 58 technologies.
- **Data security**: In seven cases, it is stated that the collected data is encrypted and separately backed up, and the rest (55) of the technologies does not provide any information on this.

3.3 First results from guided expert interviews

We conducted overall 20 Interviews (12 with inclusion experts, eight with AI experts) for a first qualitative evaluation of selected research results and seven sample technologies (one for each task group) as well as for the identification of important influencing factors and concerns. The first results of the qualitative content analysis with a specific focus on ethics and data sovereignty are presented below. Ethical issues were frequently mentioned when evaluating the suitability of sample technologies. For one technology, ethical topics were not addressed. For the remaining six technologies, positive as well as negative ethical aspects were addressed.

The following points were raised as positive:

- Enhancing participation
- Enhancing autonomy, self-determination, and independence
- Accessibility
- Increased transparency via open source
- The system adapts to people and not vice versa, e.g., via user feedback

Among the ethical concerns raised were the following:

- possibility of surveillance or extraction of information that is necessary and accepted
- neglect of other senses and thus dependency on technology, potentially leading to a lack of future investment in general accessibility
- non-acceptance of such technologies, e.g., due to permanent paternalism
- keeping the status quo: using technology to turn "silly" work into less (but still) "silly" work
- replacement of humans (reduction of staffing ratio/human assistants), risk of too much reliance on the assistance system, and resulting neglect of human interaction/reduced communication by persons with disabilities
- psychological user profiles (that may not adequately represent the user)

Topics linked to data protection and data sovereignty were rarely mentioned by the experts when evaluating the suitability of the sample technologies.

Another aim of the interviews was to identify influencing, success, or disruptive factors for transformation processes. The interviewees should imagine introducing an AI-based technology in an organization and name crucial influencing factors for a successful introduction and sustainable use of this AI-based technology. As part of the qualitative content analysis, such factors were identified and clustered into the following four groups, for which factors relating to ethical considerations were added in brackets:

- individual factors (added value in terms of job satisfaction, work facilitation)
- technological factors (no ethical aspects addressed)
- organizational factors (support of the majority of employees including management and IT: fostering a willingness to cooperate, understanding, motivation, competence)
- procedural factors (participatory approach, involve everyone from the beginning, person-centered approach)

Aspects relating to data protection appeared four times overall but were not mentioned as the first or second important influencing factors.

4 Discussion

Imagine looking for an adaptive assistive technology for a specific group of persons with disabilities that should support a particular working and learning environment. It became apparent that practitioners might find this task quite unmanageable because either essential information is not readily accessible or completely missing from the website. Also, besides legally binding aspects such as fulfilling GDPR requirements, a lack of standardization with regard to reporting basic information about each technology becomes apparent.

In the technology scanning outlined in section 3, it could be observed that information on the data that is assessed by the systems is insufficiently provided. Although data protection and GDPR conformity are among the main topics discussed in the realm of AI applications, the debate should also include concerns about persons with disabilities. One challenge is that essential information is often not barrier-free or accessible for persons with disabilities. Also, some of the AI-based assistive technologies gather data that, even when pseudonymized, could allow for the identification of a user within a user group, for instance, due to health-related data patterns. GDPR tries to protect people from being discriminated against by biased algorithmic decision-making through misused usage of sensitive data such as race, gender, or health information. Thus, even when a system is designed to assess sensor data that is not intended to be health-related, developers could ignore the fact that the assessed data may allow conclusions about the health of some target groups.

Another critical point could be that some GDPR requirements work perfectly in theory, but one might encounter challenges when looking at practical implementation. Take, for instance, the right of access or rectification. In the course of our technology scanning, it turned out to be challenging for manufacturers to provide information about the GDPR-related rights of the individual and how to exercise those in a comprehensible way. Having persons with cognitive disabilities in mind, the issue of data sovereignty is even more pressing. One means to deal with this challenge could be based on an idea by [KSL18]. Adapting and extending their idea of a self-explanation dashboard to the individual needs of users with all kinds of competence levels could be a viable solution to help users understand what data is assessed by a system, how it is processed, and with whom it is shared. It could also comprise elements that could help build up data sovereignty, explain which rights the user has and what can be done to exercise those rights. At best, this system would adapt automatically to the needs and competence level of the individual user.

In this position paper, we discussed first results of our current work on ethics-related aspects addressed in the description of AI-based assistive technologies and ethics-related opportunities and challenges stated by inclusion and AI experts. Great potential for improvement can be seen concerning aspects such as transparency and data sovereignty. When using AI-based assistance systems to support persons with disabilities, it is crucial to address the potential challenges and risks in the ethical debate to sensitize all relevant

stakeholders such as developers and manufacturers. In our future research, we will further analyze gaps in AI ethics debates and derive recommendations for further action.

Acknowledgments

This research is funded by the Federal Ministry of Labour and Social Affairs (Bundesministerium für Arbeit und Soziales, BMAS, Ausgleichsfonds für überregionale Vorhaben zur Teilhabe schwerbehinderter Menschen am Arbeitsleben) and is part of the project *KI.ASSIST – Assistenzdienste und Künstliche Intelligenz für Menschen mit Schwerbehinderung in der beruflichen Rehabilitation* (Assistance systems and artificial intelligence for persons with severe disabilities in vocational rehabilitation).

We would like to thank Anh Thu Nguyen-Hoang and Moritz Emmrich for their support in data collection and data analysis.

Bibliography

[AI19] High-Level Expert Group on AI: Ethics guidelines for trustworthy AI. European Commission, Brussels, 2019.

[Al19] Algorithm Watch, AI Ethics Guidelines Global Inventory, https://inventory.algorithmwatch.org/, accessed: 26/06/2020

[Br18] Brundage, M. et al.: The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation, 2018.

[DE19] DEK, Data Ethics Commission of the Federal Government, Opinion of the Data Ethics Commission, https://datenethikkommission.de/, accessed: 26/06/2020

[De18] Deutscher Ethikrat (German Ethics Council), Big Data and Health – Data Sovereignty as the Shaping of Informational Freedom, https://www.ethikrat.org/fileadmin/Publikationen/Stellungnahmen/englisch/opinion-big-data-and-health-summary.pdf, accessed 26.06.2020

[En18] Enquete-Kommission „Künstliche Intelligenz – Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale“ des deutschen Bundestages (Study Commission on Artificial Intelligence, German Bundestag), Website and Reports, https://www.bundestag.de/ausschuesse/weitere_gremien/enquete_ki, accessed 13 Nov 2020.

[Ha20] Hagendorff, T.: The Ethics of AI Ethics: An Evaluation of Guidelines. In Minds and Machines, 2020, 30; pp. 99–120.

[KSL18] Kiy, A.; Sass, K.; Lucke, U.: Der EU-Datenschutz-Grundverordnung begegnen: Welche Daten sind über mich erhoben und wie komme ich da ran? (Facing the General Data Privacy Regulation: What data is being collected about me and how can I get access?). In (Schiffner, D. Ed.): DeLFI-WS 2018, DeLFI Workshops 2018. Proceedings der Pre-Conference-Workshops der 16. E-Learning Fachtagung Informatik, co-located with

16th e-Learning Conference of the German Computer Society (DeLFI 2018) Frankfurt, Germany, September 10, 2018. RWTH Aachen, Aachen, Germany, 2018.

[KPN19] Kazimzade, G.; Patzer, Y.; Pinkwart, N.: Artificial Intelligence in Education Meets Inclusive Educational Technology—The Technical State-of-the-Art and Possible Directions. In (Knox, J.; Wang, Y.; Gallagher, M. Eds.): Artificial Intelligence and Inclusive Education: Speculative Futures and Emerging Practices. Springer Singapore, Singapore, 2019; pp. 61–73.

[Ma00] Mayring, P.: Qualitative Inhaltsanalyse. Grundlagen und Techniken. Deutscher Studien-Verlag, Weinheim, 2000.

[Ma13] Manzeschke, A. et al.: Ergebnisse der Studie "Ethische Fragen im Bereich Altersgerechter Assistenzsysteme". VDI, Berlin, 2013.

[OE20] OECD, Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449, accessed: 19/04/2020.

[Sa17] Sachverständigenrat für Verbraucherfragen beim Bundesministerium der Justiz und für Verbraucherschutz, Digital Sovereignty - Report by the Advisory Council for Consumer Affairs, https://www.svr-verbraucherfragen.de/wp-content/uploads/English-Version.pdf, accessed: 26/06/2020

[Yu19] Yun, H. et al.: Code of Practice for Sensor-Based Learning. In (Pinkwart, N. & Konert, J. Ed.): DELFI 2019. Gesellschaft für Informatik e.V., Bonn, 2019; pp. 199–204.

[Zi20] Ziegler, W.: A Landscape Analysis of Standardisation in the Field of Artificial Intelligence. In Journal of ICT Standardization, 2020.