How human friendly is ChatGPT for knowledge workers?
Analyzing opportunities and risks of generative AI with the FriendlyTechCheck (FTC)

Dr Anja Gerlmaier & Dr Paul-Fiete Kramer (Institute for Work, Skills and Training/University of Duisburg-Essen)

Introduction

Since the launch of ChatGPT at the end of 2022, large language models (LLMs) rapidly became a prominent technological phenomenon. The model's ability to process and generate human-like text responses makes it an innovative working tool for use in several professional settings, especially knowledge-based professions (Ali et al., 2024). Studies based on American data suggest, that occupations in areas of sales, education and research, the judiciary and administration may face more exposure to advances in generative AI (Felten et al., 2023). In this capacity generative AI like ChatGPT can provide on-demand explanations and translations. It offers guidance on various academic topics or generates source code or content.

GPT in the chatbot stands for Generative Pre-Trained Transformer. It is a dialogical language system that can anticipate linguistic patterns to questions posed with an extremely large amount of data (Schönbächler et al., 2023). Generative AI such as ChatGPT and similar systems, e.g. Gemini, Copilot or Perplexity are trained on the basis of immense amounts of data. For their use, it is important to know that the common language models encode probable word sequences in context. The results are based on linguistic probabilities of word sequences that are found in the training material. The current GPT applications are susceptible to so-called hallucinations, which includes the creation of non-real facts or quotes. In addition to this, the quality of training material has an enormous influence on content output. Discrimination or prejudices can appear in the training data (biases), and this is not necessarily recognizable for the user.

However, large language models like ChatGPT do not have a model of meaning or factual knowledge like expert systems. In so far, the usage of large language models as a working tool in knowledge-based occupations is currently the subject of controversial discussion (Mogavi et al., 2023). Some authors point out, that ChatGPT as a working tool can strengthen the productivity of knowledge workers (e.g. researchers, see Khlaif et al., 2023) and helps reducing high work load (e.g. teachers, see Ali et al., 2024). Other authors reject the use of GPT based applications in academic positions due to problems of academic integrity, risk of automation of academic tasks or overreliance on technology (Mijwil, 2023; Gmyrek et al., 2023; Hagendorff & Fabi, 2023).

A woman touching a virtual screen   © rawpixel.com - Freepik.com

In view of the described risks responsible persons are also requested to create policies and good practises for a human friendly and responsible use of AI in their organizational units. To achieve this, it is crucial to involve employers in AI adopting in an early stage: on the one hand it is easy and comparatively inexpensive for most employees to experiment with generative AI tools. Insofar, responsible persons can get crucial information about potentials and risks of AI tools from this 'early adopters'. On the other hand, it must be considered that unauthorised use of (free) AI tools can have negative outcomes for both organizations and workers because of unreliable outcomes generated by these chatbots. Research about GPT usage in knowledge-based jobs and its consequences for work quality and training is at a starting point (Gmyrek et al., 2023). For this reason, further research is necessary to identify job specified risks and opportunities of generative AI at work (Hosseini et al., 2023).

This is where our study comes in. We ask what opportunities and risks employees in the field of highly qualified knowledge work (i.e. a field of work that is particularly exposed to possible use of AI) perceive when using generative AI tools. From a socio-technical system perspective it is important to bear in mind that the use of new technologies does not necessarily lead to improvements or deteriorations of work quality and well-being (Parker & Grote, 2020). Improvements of productivity and work quality ('joint optimization') can only be reached, when new technologies fit with work flows, client's needs, workers job identification and qualification

(Winby & Mohrman, 2018; Appelbaum, 1997). Following this aim, the "HUMAINE" project started in 2021 to develop methods for human centred deployment of AI in work places. In this context, the University of Duisburg-Essen investigated in the dialogue tool 'FriendlyTechCheck'(FTC). This tool will support organizations to identify psycho-social risks and opportunities of AI based technologies at work places. Its use should be taken into account when technological changes such as the implementation of robotics or other AI based systems are planned.

In this article we report on first experiences using FTC to identify psycho-social risks and opportunities of ChatGPT in high level knowledge work. Therefore, we present the method of the FTC and its theoretical framework. Following that, we present some findings of psycho-social risks and opportunities of GPT deployment we observed in a case study with a researcher and development team. Finally, we present requirements the team pointed out for a human friendly and responsible AI usage in their research institute.

Concept of FTC framework

The "FriendlyTechCheck" (FTC) is a dialogue-oriented assessment tool to identify psycho-social potentials and risks of AI in the workplace. An aim of the assessment tool is to empower organizations to design sustainable work systems with AI (for example, in the context of risk assessments). The instrument was designed to analyze the human friendliness of new technologies, e. g. psycho-social outcomes on wellbeing, learning, autonomy, social interactions, fairness and user experience (Gerlmaier & Bendel, 2024). The checklist has its theoretical foundation in a modification of the demands-resources model (Bakker & Demerouti, 2007). The measurement includes on the one side technology related demands on action regulation, which might result in psychological or physical costs (so-called unfriendly AI). In addition to that it is possible to evaluate technological induced resources that can result from interaction with technology (so-called friendly AI). Insofar, the assessment list helps to analyze, whether an AI tool has potential to support people or is risky for personal well-being, learning, social interaction or fairness.

The procedure is designed as a checklist in which positively and negatively formulated statements about the work system can be evaluated (see table 1). So, each participant ticks the positive or negative characteristics of the system from his/her point of view. In case of perceived risks people are asked to indicate what needs to be done to minimize these risks in order to implement and use AI in a human centred way. The FTC is best carried out with the involvement of several people, which can be users, managers, technical managers or works councils. The checklist can be used for different work and technological contexts. So far, the FTC was tested for mobile and collaborative cobot applications and AI based quality control.

Table 1: Dimensions and examples of items of the FTC

Dimension	Item (example)
Wellbeing	The system can increase the user's self-esteem (e.g. reducing fear of making mistakes or minimizing time pressure).
Usability	You can use the system intuitively and correctly.
Fairness	The system means that fewer people are needed for the job.
Autonomy	There are hardly any opportunities to intervene in the system if something goes wrong.
Competence development	The system can help you familiarize yourself with new issues.
Social interaction	There may be tensions in the team that are also caused by the system (e.g. due to unclear information or different performance expectations among colleagues).

Method

In the following case study, we want to answer the question, what opportunities and risks highly skilled knowledge workers perceive when using generative AI and if the described FTC framework helps in this context to identify such positive and/or negative outcomes of ChatGPT utilization in the field of development and research jobs. The assessment list is a tool to support group discussions in organizations which plan the adoption of AI at work places. The instrument can be used in workshops or as an interview guide when less than 3 people are part of the assessment procedure (e. g. in contexts of risk analysis). The participants are first asked to fill in the total of 38 statements on a selected AI system (for example ChatGPT). In a next step the participants are asked to get into a dialogue about their reviews to risks and opportunities they identified. Afterwards, the participants get encouraged to create ideas how their organization could create a reposeful and health-oriented way of use. These ideas will be documented by the moderators and, after approval by the participants, forwarded to responsible persons for AI implementation.

Participants

The following data were obtained in a group discussion with members of a research and development team. All of them are part of a German research institute with more than 2000 members in several departments. The nine participants are mostly engineers, the age range was between 30 and 60 years. The sample consists of seven team members, one manager and the responsible person for safety and health. The majority of the participants are male. All participants already had experience in dealing with ChatGPT and related programs.

Findings

Well-being

The use of generative AI can strengthen people's health resources, for example by reducing work load or enforcing self-efficacy. Human resource damaging effects like feelings of being driven, not being needed or anxiety must be prevented. In our sample, the engineers have mixed opinions on this topic: Most of them agree, that ChatGPT can increase the self-esteem of users because it helps them to overcome cognitive impairments:

> "It's really useful if you have a writer's block, for example. That when you ask a question, something is written and you can then build on what is written to write your own texts again."

On the other side, the same number of participants think, that the system gives users the feeling to be less needed in the future:

> "But if it turns out that an artificial intelligence produces just as good results as you do, there is no reason why what you doing is still a valuable job. Then you are the universal expert for nothing at all."

The suggestions about work stress are similarly ambiguous: more than half of the team members see the opportunity to save time by using ChatGPT especially for boring administrative tasks. Others reported about harmful effects on health, e.g. feelings of permanent fear to get punishment or loss of reputation because of GPT immanent hallucinations and biases that could not brought under control by them. Furthermore, some observe addictive usage behaviour that results in long working hours and exhaustion.

> "When he (ChatGPT) then provided interesting answers, there is already the potential to continue, continue and deepen this more and more and not get out at all."

Usability

Human-centered generative AI must be designed for intuitive and barrier-free use. Additionally, it should be perceived as useful and trouble free. Most of our participants rate ChatGPT as a tool with high usability standards. The chatbot is largely intuitive to use, though more than half of them also report about low quality answers, when using it without knowledge about prompting.

> "You can apply this intuitively, but you won't get good results. You'd have to train yourself for that for example introductions, state of research or discussions you cannot write without good prompting."

All participants believe, that ChatGPT has great benefits to compensate personal weaknesses, such as language problems. But some of them also find that the accessibility of ChatGPT could be better. In addition to this, half of the participants also criticise additional efforts because of possible hallucinations which have to be brought under control.

> "Basically, I would say that ChatGPT does not bring benefits for reducing your volume of work because time savings for example, when you write a summary is redeemed by multiple checks."

Fairness

The use of generative artificial intelligence must take place with fairness aspects in mind: human-friendly AI usage should support employees and not replace them, protect humans from salary losses or digital monitoring. In our sample, the impact of ChatGPT for employment risks is evaluated inconsistently. Nearly half of the participants expect the potential of new employment chances for engineers. They suspect new job opportunities in future occupations like prompt engineers or AI system controllers. Furthermore, they believe that their jobs have a low risk of substitution because of their planning knowhow:

> "But everything that is creative, that is individual and customer-specific, that has to be negotiated, that has to be explained, where someone has an idea in their head that cannot be easily described, (...). That's where people are needed in contrast."

In contrast to these employment chances and occupational perspectives more than half of the participants fear that in future a significant part of research tasks like analyzing data, writing proposals or management summaries can be prepared in less time. In the end this could have an impact on income and employment security.

> "It's possible that the result will be that you'll no longer get full positions, but only half positions, because people will say that the work you're now doing with a full position, you can also do with half a position."

Autonomy

Interaction with generative artificial intelligence is human-friendly, if the human is in control of the process at all times and has clarity about how the system arrives at its results. In our study nearly all participants agree that ChatGPT users are always in charge of the action. The same number of them think that the chatbot supports independent working because there is less need for support from superiors or colleagues.

> "And then it's easier to ask ChatGPT than to ask a stressed colleague."

In contrast all participants share the opinion that the main problem of using ChatGPT as a work tool is the lack of controllability of the system. So, although employees feel to be in charge of the action when using the tool that does not mean that they know exactly how the system works for them. All interviewees believe that the system results cannot be tracked, which creates a feeling of loss of control. The participants report that this also leads to an increasing loss of trust in colleagues and assistants:

> "Well, but then if I'm critical and try to check the results the student has given me somewhere, I have to put in a lot, a lot of time. You can't really trust anyone anymore."

Competence development

Human-friendly AI should support users in learning without leading to long-term knowledge devaluation. In our sample, all participants see significant advantages in using GPT for knowledge building, especially when it comes to generating ideas or quickly gaining an overview of unknown issues. In addition to this, all agreed that the system could give users more chances to process high quality tasks because ChatGPT helps users to do boring tasks faster:

> "Between us, ChatGPT helps me to fill out all the forms in our administration, because the questions are so complicated that you can't answer them without Chat-GPT. And so far it's going through, but the administration doesn't know anything yet."

In the context of training and learning, all participants rate it negatively that they were not trained by their organization in the use of generative AI. They argue that prompting was a laborious and time-consuming self-study process.

> "So, if you used to do targeted group training when a new Windows or SAP was introduced and now it is common practice that further training is your individual problem. Although the company is at a disadvantage if you are not sufficiently trained for the systems because you work more slowly."

Additionally, some of the engineers see long-term negative effects of Generative AI on their level of expertise.

> "I also believe, for example, that you rely on making knowledge from ChatGPT available quickly and therefore don't store things at all. So, you don't use your memory at all, so to speak."
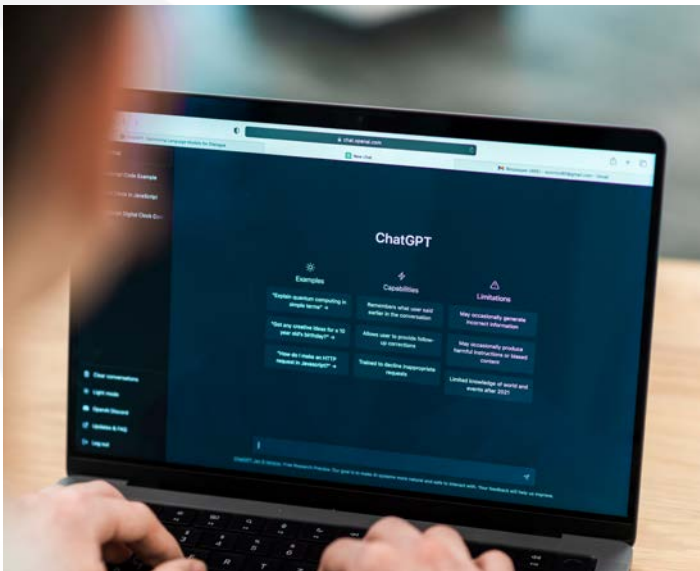
A person searching with Chat GPT on a laptop   © frimufilms - Freepik.com

Social interaction

Interaction with generative AI can change the social structures of a work system. Work should be designed in such a way that employees can interact with other colleagues or customers and make problem-solving with them. In our sample nearly half of the engineers expect more conflicts in the team when GPT is in use for research tasks. Hardly anyone sees benefits for the communication. On the contrary, six of the participants suggest generative AI as an "individualization trap": they suggest that the digital assistant could bring the risk of more social isolation and less collective problem solving in the research team:

> "I believe, for example, that ChatGPT is used relatively frequently for brainstorming. But you used to do that with colleagues and then that falls away. And that's because the others don't want to do it."

Here it becomes obvious how context-driven the outcomes of AI-usage are. While there is potential for working more independently when using ChatGPT (see 'Autonomy') these potential can turn into negative outcomes when generative AI is used in a way that undermines former collaboration-based work systems.

Users requirements for trustful and sustainable AI adoption in research and development institutes

In our second part of the FTC workshop, we asked the team members for institutional frameworks and policies to introduce generative AI in a human friendly and responsible way. Most of the group members were of the opinion that despite of the risks in using the tool for research tasks it could bring benefits for researchers' everyday life. The aim must be to promote spaces for reflective experimentation without prior prohibitions. The participants worked out two essential points for sustainable and responsible use of GPT in their research institute:

(1) Promote institutional awareness: from the team's perspective, the employer has a responsibility to sensitize employees to the potentials and risks. In their view, this is also necessary because, for example, a critical use of GPT can result in a loss of reputation for the organization:

> "That's a problem for the institution. Because then people say, yes, the people from (...), they only write fake articles."

In this context a self-learning module similar to existing formats for data protection training is proposed.

(2) Create collective training with early adopters: the participants also believe it is important to provide interested employees with learning opportunities for using GPT. In their view, this could be done, for example, through user groups in which early adopters from the workplace provide practical tips on GPT use and a regular exchange can take place.

Discussion

In this study we tried to find out, what opportunities and risks highly skilled employees in the field of research and development perceive when using ChatGPT and if the FTC framework can support organizations in identifying such outcomes of generative AI tools in knowledge-based work settings. First of all, it became obvious that participants of our study perceive both, opportunities and risks when they work with ChatGPT. However, it also became clear that these outcomes of AI usage are context-driven and depend on individual dispositions and user behaviour, organizational frameworks as well as on the specific features of the technical system. The results underline the importance of a socio-technical system perspective in order to develop and implement human centred AI. In this context, our observations in the current case with a team of engineers showed, that the FTC measure promotes a structured dialogue about friendly and responsible AI use. In the workshop, the team members were able to formulate concrete organizational needs for training and regulation. The team evaluated the measure as a good tool for a participative way of psycho-social risk management in technological change processes.

The results do not claim to be generalised. For this purpose, further investigations in other areas of knowledge work are necessary. Nevertheless, our study points to potential changes in work and employment in the field of knowledge work that seem to be worth exploring in detail.

Literature

Ali D, Fatemi Y, Boskabadi E, Nikfar M, Ugwuoke J, Ali, H (2024). ChatGPT in Teaching and Learning: A Systematic Review. Educ. Sci., 14 (6), 643. https://doi.org/10.3390/educsci14060643

Appelbaum SH (1997). Socio-technical systems theory: an intervention strategy for organizational development. MD, 35(6), 452–463.

Bakker AB, Demerouti E (2007). The Job Demands-Resources model: state of the art. JManag Psychol, 22(3), 309–328. https://doi.org/10.1108/02683940710733115

Felten EW, Raj M, Seamans R (2023). Occupational Heterogeneity in Exposure to Generative AI. http://dx.doi.org/10.2139/ssrn.4414065

Gerlmaier A, Bendel A (2024). Humanzentrierte Bewertung und Gestaltung von autonomen Transportsystemen: ein Fallbeispiel aus der Distributionslogistik. Beitrag zum 70. GfA-Frühjahrskongress – "Arbeitswissenschaft in-the-loop".

Gmyrek P, Berg J, Bescond D (2023). Generative AI and jobs: A global analysis of potential effects on job quantity and quality. ILO Working Paper 96. https://ssrn.com/abstract=4584219

Hagendorff, T, & Fabi, S (2023). Why we need biased AI: How including cognitive biases can enhance AI systems. Journal of Experimental & Theoretical Artificial Intelligence, 1–14. https://doi.org/10.1080/0952813X.2023.2178517

Hosseini M, Gao CA, Liebovitz DM, Carvalho AM, Ahmad FS, Luo Y, MacDonald N, Holmes KL, Kho A (2023). An exploratory survey about using ChatGPT in education, healthcare, and research. PLoS ONE, 18(10). https://doi.org/10.1371/journal.pone.0292216

Khlaif ZN, Mousa A, Hattab MK, Itmazi J, Hassan AA, Sanmugam M, Ayyoub A (2023). The Potential and Concerns of Using AI in Scientific Research: ChatGPT Performance Evaluation. JMIR Med Educ, 2023 Sep 14. https://doi.org/10.2196/47049

Mijwil MM, Hiran KK, Doshi R, Dadhich M, Al-Mistarehi, A-H, Bala I (2023). ChatGPT and the future of academic integrity in the artificial intelligence era: a new frontier. Al-Salam Journal for Engineering and Technology, 2(2), 116-127.

Mogavi RH, Deng C., Kim JJ, Zhou P, Kwon YD, Metwally AHS, Tlili A, Bassanelli S, Bucchiarone A, Gujar S, Nacke LE, Hui P (2023). Exploring user perspectives on ChatGPT: Applications, perceptions, and implications for AI-integrated education. https://doi.org/10.48550/ARXIV.2305.13114

Parker SK, Grote G (2020). Automation, Algorithms, and Beyond: Why Work Design Matters More than Ever in a digital world. Applied Psychology, 71(4), 1169-1653. https://doi.org/10.1111/apps.12241

Schönbächler E, Strasser T, Himpsl-Gutermann K (2023). Vom Chat zum Check. Informationskompetenz mit ChatGPT steigern. Medienimpulse, 61(1). https://doi.org/10.21243/mi-01-23-18

Winby S, Mohrman SA (2018). Digital Sociotechnical System Design. The Journal of Applied Behavioral Science, 54(4), 399-423. https://doi.org/10.1177/0021886318781581