European Agency for Safety and Health at Work

Cognitive automation: implications for occupational safety and health

Executive Summary

European Agency for Safety and Health at Work

Safety and health at work is everyone's concern. It's good for you. It's good for business.

Authors: Patricia Helen Rosen, Federal Institute for Occupational Safety and Health (BAuA), Robert Donoghue, University of Leicester, School of Business, Eva Heinold, Federal Institute for Occupational Safety and Health (BAuA), Prof. Dr Phoebe Moore, University of Leicester, School of Business, Susanne Niehaus, Federal Institute for Occupational Safety and Health (BAuA), Dr Sascha Wischniewski, Federal Institute for Occupational Safety and Health (BAuA)

This is a summary of a report that was commissioned by the European Agency for Safety and Health at Work (EU-OSHA). Its contents, including any opinions and/or conclusions expressed, are those of the authors alone and do not necessarily reflect the views of EU-OSHA.

Project management: Ioannis Anyfantis, Annick Starren, Emmanuelle Brun (EU-OSHA).

1 Introduction

As a consequence of digitalisation, jobs and working tasks are continuously changing. The development of recent technologies, such as artificial intelligence (AI) and advanced robotics, has especially established new possibilities for task automation and revived the debate on work-related psychosocial and organisational aspects and on workers' safety and health. This report explores the implications for occupational safety and health (OSH) associated with AI-based systems and cognitive automation. As the report demonstrates, the effects are far-reaching. A wide range of different kinds of cognitive tasks are already being automated, and this trend is set to continue at great speed. The report serves as a key resource for policy-makers by providing a thorough analysis of the risks and opportunities associated with AI and cognitive automation. The critical points are summarised here.

Given the relatively narrow state of AI today, it makes more sense to discuss the automation of tasks as opposed to entire occupations. Therefore, a three-part taxonomy is introduced to help organise the findings. These task categories are differentiated by that which a labourer acts on in the production process. Person-related tasks involve a worker interfacing with a person (customer, patient), information-related tasks with information (data processing, software programming, etc.), and object-related tasks with objects (vehicles, etc.). Whilst AI will automate the completion of all these task types, the literature demonstrates that, for now, information-related tasks are the most suited for automation by AI-based systems, when referring to cognitive tasks.

2 Impact on working environment and occupational safety and health implications

Since the technologies responsible for cognitive automation like algorithmic programming, AI and so on are of general purpose, they can have a far-reaching impact in a very short period of time. Such technologies do not have sector-specific limitations and can be applied across industrial sectors simultaneously and with great rapidity. The literature surveyed for the report offers empirical verification of this point. The ever-growing capability of AI to complete any person-, information- and object-related tasks is already measurably transforming the educational, medical, legal, financial and public sectors, among others.

The report details many of these kinds of tasks that are increasingly completable by AI systems. They include, but are not limited to, customer support, well-being management, customised teaching, classroom assessment and supervision, health monitoring, decision-making and diagnoses, personal financial advice and data classification. For now, AI-based systems are able to execute a large number of information-related tasks. There is also evidence of a rapidly escalating capacity to complete person-related tasks, especially in the care industry that is facing a crisis with ageing demographics. With respect to object-related tasks, AI is considerably less impactful. The report singles out autonomous driving vehicles as the upcoming major advancement related to this category of task automation, but it is not yet clear to what extent this will impact work within transportation or delivery services.

Therewith, it presents opportunities and challenges for OSH associated with the automation of cognitive tasks. In addition the report addresses cybersecurity as a topic that need to be addressed on an organisational and legislative level to ensure OSH at the workplace. AI-based systems might find themselves as both target and executing force of cyberattacks, as their capabilities increases, putting personal data at risk. However, it can also play a key role in protecting said data (Oancea, 2015).

Tools for risk assessment provide a basis for OSH oriented decision making. However, as AI-based systems are an emerging technology, there currently is a lack of tools, legislation or guidelines assisting companies in risk identification and risk analysis. The European Commission provides risk categorisations for AI, and more applicable regulation is said to be published in 2024 (European Commission, 2021). Until then and beyond other tools are needed which facilitate risk assessment of AI-based systems in the workplace to ensure OSH.

2.1 Opportunities for improved working conditions

The automation of cognitive tasks by AI-based systems will continue to eliminate repetitive and boring kinds of clerical or administrative work. As intelligent programs more efficiently process forms, applications, claims, legal documents and so on, it will no longer be necessary for humans to complete these 'mind-numbing' and alienating tasks. If systems analyses and recommendations of AI systems prove to be effective and accurate enough to be considered worthy of trust and more broadly followed, administrators could potentially either supervise more projects or focus more on the human-centred side of their job. This would constitute a shift to potentially more engaging kinds of work and reduce the cognitive workload.

Physical object-related tasks are likely the most well-known form of application for robotic systems. Technological progress is not a neutral phenomenon: it raises the prospect of both positive and negative developments. Many are confident that advancements in AI will continue the historical trend of eliminating dangerous jobs. The most comprehensive example of how the automation of a cognitive task can have physical implications would be the advent of self-driving vehicles. Approximately 9.3 individuals per 100,000 die each year in traffic-related fatalities in Europe. A considerable proportion of people on the road at any given time are commuters driving to work, ride service providers, or truck drivers transporting goods and services. It is widely believed that the rise of self-driving vehicles could dramatically minimise this cause of premature death. While driving has prominent physical components, the inbuilt AI-based systems primarily automate perception-based tasks of a driver and based on their analysis trigger the appropriate physical response (e.g. braking) in the vehicle. Investments in life-saving technologies have a great deal of potential upside both in terms of preventing premature and needless deaths and limiting healthcare costs associated with accidents. Another related hope is that AI can reduce the burdensome and emotionally taxing nature of some occupations. Care work, for example, is currently a very high-touch occupation. That is, carers have to constantly engage in physical and emotional interaction with patients when completing the entirety of their job duties. If some aspects of providing care can be offloaded onto smart devices, this could transform care work into an increasingly low-touch labour process, and thereby curb the emotionally challenging dimensions of the work as it is performed now.

2.2 Impact on sectors

The analysis of automated cognitive tasks among sectors reveals a high number of automated or supported tasks in the sector of **human health and social work activities**. Here, the majority of tasks can be found in hospital activities. The plethora of possible applications for AI-based systems indicates that in the near future the installation of such technology in this working environment will gain momentum. On a sectoral level, healthcare and social work is likely to continue to grow in its importance and also as a major field of application for AI-based systems. Secondly an extensive body of scientific literature is dedicated to the **education** sector. Educational platforms that incorporate AI based components or AI based software solutions are currently used to assist those working in the sector, offering capabilities that can improve teaching procedures and reduce cognitive workload. Furthermore, the general sector of **Professional, Scientific and Technical Activities** as well as **Administrative and Support Service Activities** sectors are also addressed quite frequently in scientific literature and mentioned by the experts, given the broad distribution of automated software systems.

2.3 Occupational safety and health concerns

The majority of occupational safety and health (OSH) implications arising when AI systems are used for the automation of cognitive tasks lie within the psychosocial realm. As this report focuses on the automation of cognitive tasks, this result is not entirely surprising. Major risks that are listed independently of any given sector, job or task are the fear of job loss, negative impacts of job transformations, mismatched trust in the system and the possible loss of autonomy through it. In addition to that, while discussed most prominently in the area of teaching, the loss of privacy is a noticeable concern that can be applied to the more general usage of AI-based systems. The potential for increased loss of privacy is in particular different from previous automation fears, because AI-based systems by design often gather and to some extent analyse data. For ethical reasons, workers need to be aware if this is happening, and if it is, what data are being collected and what this data is used for. Furthermore, any AI-based system in the workplace that collects data should abide by the most recent ethics and privacy and data protection regulations. While the fear of job loss is a psychosocial experience and therefore can be considered ‘subjective’, the actual risk of task replacement and thus aspects of job loss because of the introduction of AI-based systems are not. However, there is no consensus among experts as to the actual extent of it, where there is an imbalanced ratio between jobs destroyed and jobs created, in this climate.

2.3.1 Job loss

The most obvious concern is the threat of job loss. There have been many studies in recent years attempting to 'calculate' how many jobs will be made redundant within a given time horizon, which is not an entirely appropriate question. It is nevertheless the case that large numbers of workers currently believe that their job will be automated within the next few years. This is problematic given the repeatedly confirmed finding that there is a strong relationship between feelings of precariousness at work and poor mental health.

2.3.2 Job transformation

Because the roll-out of AI is more likely to eliminate tasks as opposed to entire jobs and occupations, there will be widespread and continual occurrences of job transformation. Deskilling is a serious risk associated with continuously changing job content. As certain skills become less demanded in the labour market and necessary in particular occupations, people who have those skills will likely lose the ability to perform them over time. Of particular concern is the prospect of moral deskilling. As algorithms are deployed to replace humans in making decisions that have moral content, individuals' capacity to perform moral reasoning may atrophy. Deskilling, of all kinds, is likely to have a corrosive effect on society.

The touted solution to the problem of deskilling, what is sometimes called 'upskilling' or 'reskilling', also presents OSH risks. First, it is not clear that it actually yields the assumed results. Kunst's analysis concluded that 'while increasing human capital investments may be necessary, they do not guarantee success on the labour market: in spite of the substantial skills that they had acquired, manufacturing craftsmen have experienced pervasive declines in relative wages and employment opportunities since the 1950s' (2019, p. 28). Second, the pressure to upskill can amount to an oppressive burden that leads to rising stress levels. This is particularly true with more advanced AI systems. Surya (2019) explains that increased uptake of AI would 'radically revise a certain kind of training required during the next era' (p. 9). As the author point out, it 'is challenging to acquire the requisite skills to implement AI technological innovations,' and therefore workers may not 'feel confident interacting with technology or be aware of current regulations, like privacy and data legislation that directly impact AI ventures'.

2.3.3 Trust

While the possible gravity of insufficient trust or unregulated automation bias can vary from workplace to workplace, it is advised to always consider it. A general takeaway is that for any user to fully benefit from the system, they need a sufficient level of trust towards it. This can result in direct effects, like fully benefiting from the system's intended effect of cognitive support, to more indirect effects by avoiding the consequences of automation bias, in the form of over reliance or loss of skill. When introducing a new system to a workplace, everyone in contact with it should be made aware of the capabilities and realistic limitations of the system. Users should be given training to not only understand the technology but also see how their work changes due to it.

2.3.4 Loss of autonomy

Autonomy is regarded as a constituent feature of meaningful work, and, therefore, encouraging its preservation and expansion should be a goal of policy-makers where appropriate. In this respect, the dispersion of AI into workplaces presents complications and challenges. First and foremost, new technologies can have a constricting effect on the totality of the work execution process. Smids et al. (2020) explain that 'some robotic applications in the workplace may require working according to a very strict protocol that leaves little room for human creativity, judgment, and decision-making. For the same reasons, workers' opportunities to engage in job crafting may be severely restricted' (p. 514). In short, restricted choice in the execution of one's work entails that one's 'autonomy would be undermined, and consequently the jobs' meaningfulness as well' (p. 514).

2.3.5 Privacy

Loss of privacy is another central concern related to the deployment of AI in workplaces. Widespread data collection is required for AI systems to operate. Thus, the implementation of such systems involves numerous and complex questions regarding consent, selection, transparency, representation and accountability, among other considerations that arise when a population is monitored and their data are collected (Köbis & Mehner, 2021). Failure to develop and enforce ethics guidelines for the collection and utilisation of instruction-related data could result in widespread rights violations for educators.

Another negative association between monitoring and workplace freedom has to do with the phenomenon of self-censorship. When individuals are aware that they are being watched, they may feel an innately arising pressure to act in what they believe is the most desirable manner in the eyes of the observer. An employee under constant monitoring may believe they must work with greater intensity than they actually have to, feeling that if they are observed moving at the wrong pace they could be disciplined. In this sense, they have suffered a loss of freedom to exercise basic workplace rights like working to their actual contracted duties.

2.3.6 Depersonalisation

The literature surveyed for this report, particularly findings from the care and education industries, suggests that the uptake of AI could induce a process of depersonalisation. The introduction of AI into the care industry is uniquely illustrative. Rubeis (2020) explains that the expansion of smart 'technology leads to the distinction between patients as bodies and patients as subjects' (p. 2) because the central focal point of care becomes 'easily measurable indicators that are usually bodily in nature' (p. 2). In other words, growing involvement of monitoring systems and instructional assistants in the caregiving process transform the relationship between the carer and patient, ultimately by turning the latter into an object for the former. The patient no longer represents their needs as a subject, rather, their needs are directly observed by the carer through technological devices.

Although the literature tends to focus on the potential benefits and harms of AI technologies for patients, we can reasonably assume that depersonalisation in the nursing relationship may promote a form of alienation for caregivers. As more aspects of care work become automated, the care worker's responsibilities are revolutionised from actively assessing patient needs and prescribing a course of action to responding to alerts and following machine-generated recommendations. This reconfiguration from active assessment and prescription to following mechanical commands alienates and limits the projection of the carer into their work. Put another way, the workers no longer extend themselves into the decision-making processes, effectively limiting the need to utilise their emotional and cognitive capacities whilst providing care. Another example is workers in the delivery sector, who previously were able to actively plan their route but now have to follow an algorithmically optimised route.

Another related concern is the dehumanisation of an increasingly automated work environment. As more tasks are offloaded onto computer systems, all types of robots, instructional assistive technologies and so on, care workers are increasingly surrounded by, and reacting to, 'data' and 'devices' more than interacting with human beings. For those who enter this line of work because they value the socially interactive element of caring for others, this will become a less central feature of care work, thus depriving them of that opportunity. Such deprivation amounts to a harm as it effectively blocks an individual from participating in an activity linked to their own self-actualisation and fulfilment from work.

2.3.7 Cybersecurity

The topic of cybersecurity needs also to be addressed on an organisational and legislative level to ensure OSH at the workplace when using AI-based systems. Especially, if the AI-based system is handling sensitive data, such as personal data, or in case of cobots if the system is interacting directly with a worker. Where and how to use AI in the context of cybersecurity, including how to protect an AI-based system and the data it processes form potential outside intervention, is an organisational consideration companies will likely phase in the future.

2.4 Occupational safety and health benefits

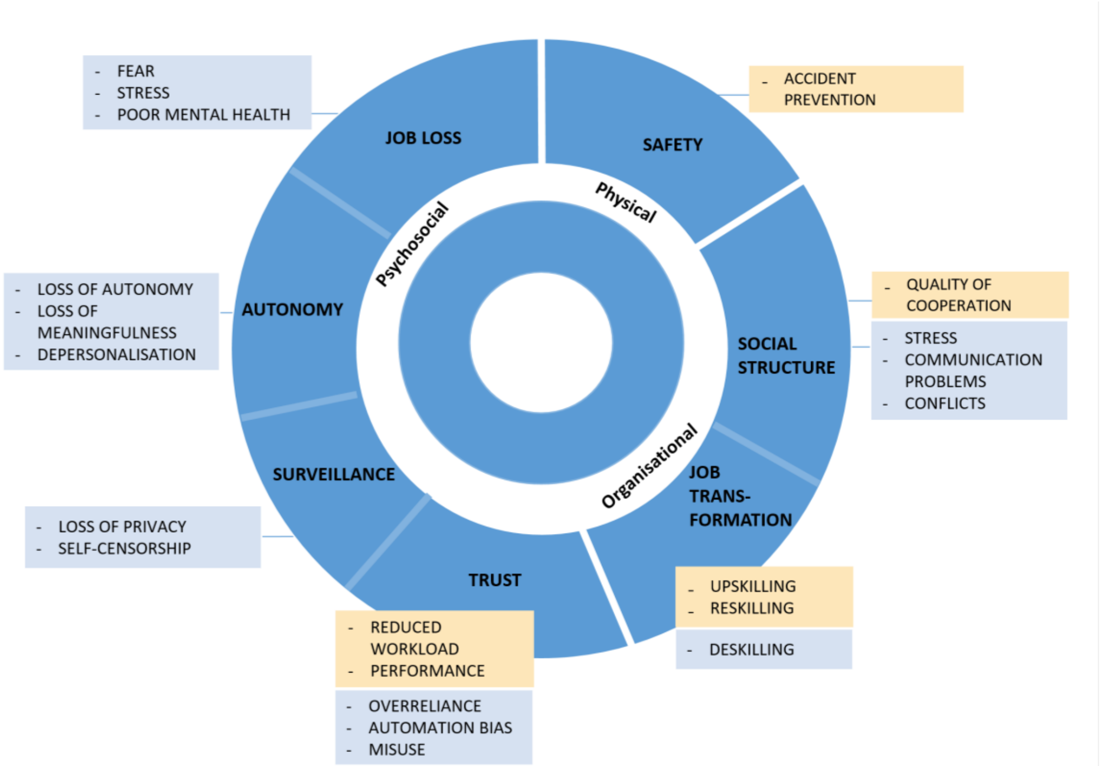
Next to the risks of applying AI-based systems in a workplace, there can also be a variety of benefits. Similar to the risks, they often fall under psychosocial benefits. The elevation of mental workload and stress are by far the most prevalent benefits discussed. However, the actual effect that automating a task has on the mental capacities of a worker is often not researched in depth; while mental workload and stress are two aspects of psychosocial effects caused by the system, the longevity of the effect should be investigated in greater depth. As workers might get used to the new workload, they may fill their capacities with new tasks that arise from using the AI system. While some positive impact of AI-based systems on physical OSH have been mentioned in literature, they were mostly a peripheral effect. Examples of this are improved safety surveillance systems and decision support systems that support a worker during a crisis situation. The most common tangible physical OSH benefit is through AI being used to reduce traffic accidents, as a result of automating a cognitive task. While not directly related to a specific job, using it can be potentially live saving for a significant amount of workers. Effective training of workers could also benefit from such systems.

Nonetheless, in the education sector, monitoring processes may allow for helpful feedback, student customisation potential, time saving and so on. Greater surveillance opens up the possibility for the collection of incriminating information as well – information that could be used to bring forward more frequent disciplinary sanctions against poor performance. In this sense, the site of educational instruction would, like other highly digitalised spaces, become increasingly panoptic. The rising rate of teacher observations as a means to improve education outcomes demonstrates a tolerance and willingness for classroom monitoring, something that AI could take to whole new levels.

Furthermore, suitable AI-based systems such as a decision support systems can be useful to help mediate communication and coordination problems in modern complex corporate environments, reducing stress.

Overall, there are a number of complex risks and benefits related to AI-based systems in the workplace. They prominently fall into the categories of psychosocial and organisational risks for workers, often not specifically bound to a singular job or sector. Figure 1 provides an overview of both risks and opportunities resulting from the integration of AI-based systems in the workplace for the automation of cognitive tasks.

Figure 1: Overview of OSH relevant factors and effects



2.5 Risk assessment

The specific OSH impact of introducing an AI-based system into a workplace is often hard to gage and varies dependent on the specific system, automated task and environment. So is the overall risk assessment of implementing an AI-based system into the workplace. While there are AI-based tools to perform risk assessment on specific areas of application, like medical conditions, risk assessment tools for AI-based systems and their impact on OSH are currently an area which lacks in options. Accurate and in-depth risk assessment of a technology in the workplace is vital to ensure OSH, and the lack of assessment tool capable of providing this for AI-based systems needs to be considered going forward.

3 Conclusion

The report highlights the opportunities and risks of AI and cognitive automation to provide policy-makers with knowledge and understanding of the state of play within this arena, around current and forthcoming smart technologies. A primary objective of policy-makers should be to encourage the opportunities that improve social welfare and dampen possible harmful consequences. The report identifies a number of key risks that should be addressed by policy-makers through better enforced and revised labour laws or data protection regulations (also see Moore, 2020).

Of course, the field of AI-based systems in the workplace is diverse and rich in detail. It is possible to categorise some of it along the lines of degree of automation, task category, and identified challenges and opportunities based on the current state of research on AI-based systems in the workplace. However, it is equally important to acknowledge the complexity within each system that is unique to its application. As the automation of cognitive tasks progresses with rapid speed, researchers and policy-makers need to focus on OSH-relevant topics, while also addressing the current gaps in research, to ensure a human-centred, or 'human in command' approach to the development and integration of AI-based systems in the workplace.

References

European Commission (2021). Regulatory framework proposal on artificial intelligence. European Union. https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

Köbis, L., & Mehner, C. (2021). Ethical questions raised by AI-supported mentoring in higher education. *Frontiers in Artificial Intelligence, 4*. doi:10.3389/frai.2021.624050

Kunst, D. (2020). Deskilling among manufacturing production workers (SSRN Scholarly Paper ID 3429711). *Social Science Research Network*. doi:10.2139/ssrn.3429711

Kong, P., Li, L., Gao, J., Liu, K., Bissyandé, T. F., & Klein, J. (2018). Automated testing of android apps: A systematic literature review. *IEEE Transactions on Reliability, 68*(1), 45-66. doi:10.1109/TR.2018.2865733

Manokha, I. (2018). Surveillance, panopticism, and self-discipline in the digital age. *Surveillance & Society, 16*(2), 219-237. doi:10.24908/ss.v16i2.8346

Moore, P. V. (2020). *Data subjects, digital surveillance, AI and the future of work*. Scientific Foresight Unit, European Parliamentary Research Service. https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU(2020)656305

Neumerski, C. M., Grissom, J. A., Goldring, E., Drake, T. A., Rubin, M., Cannata, M., & Schuermann, P. (2018). Restructuring instructional leadership: How multiple-measure teacher evaluation systems are redefining the role of the school principal. *The Elementary School Journal, 119*(2), 270-297, doi:10.1086/700597

Oancea, C. (2015). Artificial Intelligence Role in Cybersecurity Infrastructures. International Journal of Information Security and Cybercrime, 4 (1), 59-62. doi: 10.19107/IJISC.2015.01.08

Rubeis, G. (2020). The disruptive power of artificial intelligence. Ethical aspects of gerontechnology in elderly care. *Archives of Gerontology and Geriatrics, 91*, Article 104186. doi:10.1016/j.archger.2020.104186

Smids, J., Nyholm, S., & Berkers, H. (2020). Robots in the workplace: A threat to—or opportunity for—meaningful work? *Philosophy & Technology, 33*(3), 503-522. doi:10.1007/s13347-019-00377-4

Surya, L. (2019). Artificial intelligence in public sector. *International Journal of Innovations in Engineering Research and Technology, 6*(8), 7-12. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3785663

The European Agency for Safety and Health at Work (EU-OSHA) contributes to making Europe a safer, healthier and more productive place to work. The Agency researches, develops, and distributes reliable, balanced, and impartial safety and health information and organises pan-European awareness raising campaigns. Set up by the European Union in 1994 and based in Bilbao, Spain, the Agency brings together representatives from the European Commission, Member State governments, employers' and workers' organisations, as well as leading experts in each of the EU Member States and beyond.

European Agency for Safety and Health at Work

Santiago de Compostela 12
48003 - Bilbao, Spain

E-mail: information@osha.europa.eu

https://osha.europa.eu

Publications Office
of the European Union