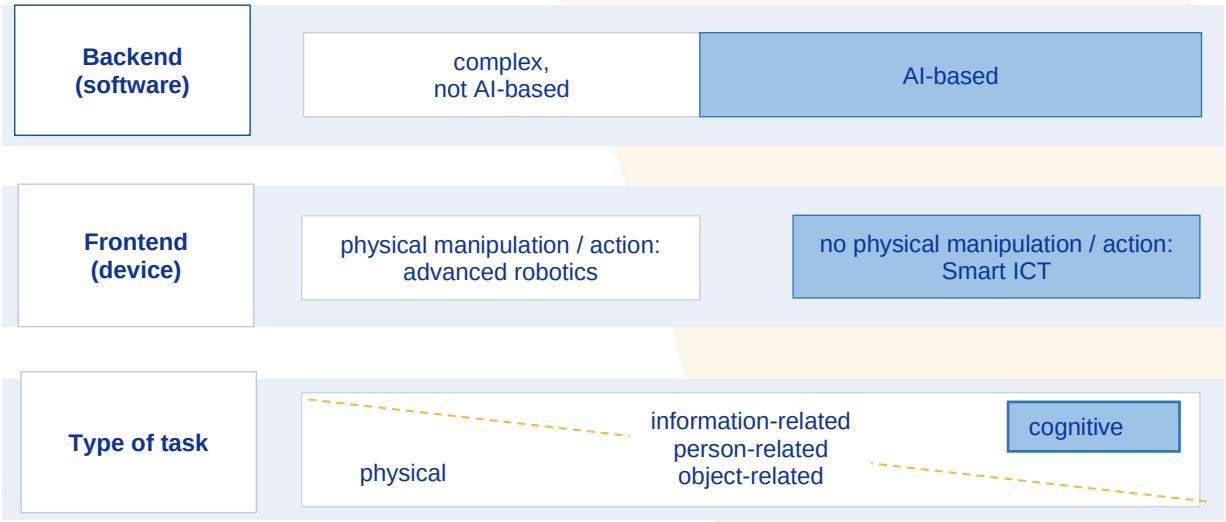
COGNITIVE AUTOMATION: IMPACT, RISKS AND OPPORTUNITIES FOR OCCUPATIONAL SAFETY AND HEALTH

AI and Cognitive Automation

According to a recent proposal by the EU Commission, 'artificial intelligence system (AI system)' means 'software that is developed with one or more of the techniques and approaches […] and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with'[1]. We see more and more automation of cognitive tasks through AI-based systems in the workplace, some more prevalent and others more subtle. When applied in workplaces, AI-based systems are mostly narrow in function and lack general intelligence. This means that their incorporation into workplaces is likely to make particular tasks of an occupation redundant for humans to perform, rather than eliminate entire occupations outright, as they need critical thinking and decision making on a level of complexity not yet achievable by AI.

Figure 1: Extract of the taxonomy for AI-based systems and advanced robotics for the automation of tasks with categories relevant for cognitive task automation



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, drones, etc.). While AI-based systems are able to automate tasks of all three types, the literature demonstrates that, for now, information-related tasks are the most suited for automation by AI-based systems, followed by person-related tasks.

Task automation

Person-related tasks are rapidly being automated by smart software. Two interaction systems are now predominantly being deployed to replace traditionally person-to-person interactions: a) chatbots, and b) AI-based conversation agents. The former refers to a system that uses natural language processing in written form, and the latter in spoken form to interact with someone. These can be applied in a number

[1] EU Commission (2021). *Laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts*.

of work settings such as telephone-based customer support[2,3] or for patients' wellbeing management through conversation-based health monitoring[4]. In customer support, these systems might perform a telephone conversation with a customer where they are either able to resolve the presented issue or redirect the customer to a specialised human operator. The same can be said for chatbots in a digital environment.

Numerous examples of information-related tasks being automated can be found in the literature, as well as in actual workplaces today. One that receives considerable attention is health monitoring. Health monitoring of patients is a crucial part of medical procedures, which is routinely performed by medical staff. The process often includes standardised health assessments to track a patient's condition, identify behaviour changes, or screen for possible complications. Necessary parts of this data are either self-reported by the patient to a healthcare professional or based on a dialogue between the two parties. A variety of conversational agents using natural language processing and free text input have been developed to address these specific tasks, as their standardised nature allows for efficient automation. These types of conversational agents focus on data collection and processing[5]. The collection itself can potentially be fully automated, however, many of these systems are still in their developing stages[6].

Finally, considerably fewer examples can be found of AI-based systems automating object-related tasks. They mostly focuses on the operation of vehicles and are currently oriented toward supporting rather than replacing drivers. Within the task of driving, we find several ways an AI-based systems can be introduced to assist drivers with their tasks. These can be, but are not limited to, smart intersection warning and rear-end warning, lane departure, driving takeovers, or cruise control with the specific focus on collision prevention[7,8,9]. Researchers have been able to provide recommendations for which specific driving task, such as lane switching, should receive prioritised support by an AI-based system. Introducing AI-based systems to support drivers significantly raises safety effectiveness because these systems work to reduce human errors[9].

Impact on jobs and sectors

Many tasks automated by AI-based systems are applicable cross-sectoral and not unique to just one specific job.

The growing capacities of AI-based systems to complete each task type will probably impact a wide range of jobs and industrial sectors in the near- and long-term future. The **medical and healthcare industry** is frequently described as being on the brink of major transformation due to smart technologies. We find that data-based processes in the medical field are being automated, while higher cognitive tasks, like the final diagnosis or treatment plan, are still made by skilled medical professionals. However, as the technology becomes more advanced, it is within reason that their assessment becomes less supervised. Some medical

[2] Bavaresco, R., Silveira, D., Reis, E., Barbosa, J., Righi, R., Costa, C., Antunes, R., Gomes, M., Gatti, C., Vanzin, M., Júnior, S.C., Silva, E., & Moreira, C. (2020). Conversational agents in business: A systematic literature review and future research directions. *Computer Science Review*, *36*, 100239.

[3] Tuomi, A., Tussyadiah, I., Ling, E. C., Miller, G., & Lee, G. (2020). x=(tourism_work) y=(sdg8) while y= true: automate (x). *Annals of Tourism Research*, *84*, 102978.

[4] Federici, S., de Filippis, M. L., Mele, M. L., Borsci, S., Bracalenti, M., Gaudino, G., Cocco, A., Amendola, M., & Simonetti, E. (2020). Inside pandora's box: a systematic review of the assessment of the perceived quality of chatbots for people with disabilities or special needs. *Disability and rehabilitation: Assistive technology*, *15*(7), 832-837.

[5] Rheu, M., Shin, J. Y., Peng, W., & Huh-Yoo, J. (2021). Systematic review: trust-building factors and implications for conversational agent design. *International Journal of Human-Computer Interaction*, *37*(1), 81-96.

[6] Milne-Ives, M., de Cock, C., Lim, E., Shehadeh, M. H., de Pennington, N., Mole, G., & Meinert, E. (2020). The effectiveness of artificial intelligence conversational agents in health care: Systematic review. *Journal of Medical Internet Research*, *22*(10), e20346.

[7] De Winter, J. C., Happee, R., Martens, M. H., & Stanton, N. A. (2014). Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence. *Transportation Research Part F: Traffic Psychology and Behaviour*, *27*, 196-217.

[8] McDonald, A. D., Alambeigi, H., Engström, J., Markkula, G., Vogelpohl, T., Dunne, J., & Yuma, N. (2019). Toward computational simulations of behavior during automated driving takeovers: A review of the empirical and modeling literatures. *Human Factors*, *61*(4), 642-688.

[9] Wang, L., Zhong, H., Ma, W., Abdel-Aty, M., & Park, J. (2020). How many crashes can connected vehicle and automated vehicle technologies prevent: A meta-analysis. *Accident Analysis & Prevention*, *136*.

devices (for example, blood pressure monitors) already include software that assesses the patient's count accurately enough that human reassessment is only necessary in exceptional situations[10].

Another sector facing considerable disruption is **teaching and education**. Advancements in AI-based systems have vastly expanded the scope of what can be performed by non-human agents in this field. A critical innovation is the development of Intelligent Tutoring Systems (ITSs). Hernández de Menéndez and colleagues, for instance, describe how these ITSs are capable of 'helping students in their daily educational activities by interpreting their responses and learning as they operate. The algorithms can offer the student problems to solve or specific videos based on his past or current interactions'[11]. Advanced versions can even provide personalised learning experiences and customised material developed from constant student monitoring. As these kinds of platforms are rolled out into the education system, teaching will be considerably altered.

In a broader job category, we see that **administrative positions** can receive support from AI-based systems too. Here the AI, a prominent candidate being decision support systems, pre-evaluates input about the current state of the situation or project, and based on this data suggests a course of action or identifies the next planning steps[12,13,14]. This kind of technology could be applied in almost **any work context** and job that requires planning and coordination. This could have major implications for industries that engage in voluminous processing of forms and applications like public services or insurance providers, but also in fields that are not primarily associated with these tasks, like the construction industry.

Artistic and creative work is usually considered non-routine and is not seen to be in great danger of task replacement. According to some authors[15], AI systems can achieve a minimal degree of creativity by copying and mimicking human ideas and rules existing in original works; however, imagination and creativity are human mental processes of high complexity. Therefore, AI can only generate artistic work with human intervention.

Occupational safety and health (OSH) considerations

Technological advancement is often a double-edged sword in that it presents both risks and opportunities. Many are hopeful that advancements in AI-based systems will continue the historical trend of eliminating dangerous jobs. A prominent example of this 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.

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, etc., it will no longer be necessary for humans to complete these 'mind-numbing' and alienating tasks. If analyses and recommendations made by 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.

Another related hope is that AI-based systems can reduce the burdensome and emotionally-demanding 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

[10] Pappaccogli, M., Di Monaco, S., Perlo, E., Burrello, J., D'Ascenzo, F., Veglio, F., Monticone, S., & Rabbia, F. (2019). Comparison of automated office blood pressure with office and out-off-office measurement techniques: A systematic review and meta-analysis. *Hypertension*, *73*(2), 481-490.

[11] Hernández de Menéndez, M., Escobar, C., & Morales-Menendez, R. (2020). Technologies for the future of learning: State of the art. *International Journal on Interactive Design and Manufacturing*, *14*. https://doi.org/10.1007/s12008-019-00640-0

[12] Garousi, V., & Mäntylä, M. V. (2016). When and what to automate in software testing? A multi-vocal literature review. *Information and Software Technology*, *76*, 92-117.

[13] Desgagné-Lebeuf, A., Lehoux, N., & Beauregard, R. (2020). Scheduling tools for the construction industry: Overview and decision support system for tool selection. *International Journal of Construction Management*, 1-12

[14] Aslam, A., Ahmad, N., Saba, T., Almazyad, A. S., Rehman, A., Anjum, A., & Khan, A. (2017). Decision support system for risk assessment and management strategies in distributed software development. *IEEE Access*, *5*, 20349-20373.

[15] Gudkov, A. (2020). Robot on the shoulders of humans. *The Journal of World Intellectual Property*, *23*(5-6), 759-776.

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.

Considerations on an organisational and legislative level should also go towards the topic of cybersecurity. AI will increasingly handle threats in that area[16] by taking over tasks like high accuracy threat detection with higher efficiency compared to human intervention[17], effectively supporting IT workers and cybersecurity specialists. However, it also can become a target of such attacks. The OSH related consequences of this can vary widely. The risk of a system being targeted needs to be considered especially if the AI is handling sensitive or personal data. However, it can also play a key role in protecting said data[16].

Additionally there is a lack of tools supporting companies in conducting a thorough risk assessment of AI-based systems when it comes to OSH. While there are some publication addressing emerging risks associated with AI in the workplace, like EU-OSHA policy brief „Impact of artificial intelligence on occupational safety and health“[18], and the European Commission regulatory framework[19] on artificial intelligence, which defines four levels of risk in AI applications, these provide a first overview. Specific tools developed to assess and AI and its surroundings to highlight possible risks specific to this set-up could support companies in ensuring OSH.

OSH risks

EU-OSHAs report on **‘Artificial Intelligence and automation of cognitive tasks**: **Implications for occupational safety and health’** identified a number of key risks that should be addressed by policy-makers by looking at labour law and data protection regulation. 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. As noted above, this isn’t an entirely appropriate question. However, 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 precarity at work and poor mental health.

Because the roll out of AI-based systems 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 needed in the labour market and 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 morally reason may atrophy. Deskilling, of all kinds, is likely to have a corrosive effect on society.

The touted solution to the problem of deskilling, sometimes referred to as **‘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’[20]. 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-based systems. Surya explains that increased uptake of AI-based systems would ‘radically revise a certain kind of training required during the next era’[21]. As pointed out, ‘[it] is challenging to acquire the requisite skills to implement AI-based technological innovations,' and therefore workers may not ‘feel confident interacting with

[16] Oancea, C. (2015). Artificial Intelligence Role in Cybersecurity Infrastructures. International Journal of Information Security and Cybercrime, 4 (1), 59-62. https://doi.org/10.19107/IJISC.2015.01.08
[17] Tschider, C. (2018). Regulating the IoT: Discrimination, Privacy, and Cybersecurity in the Artificial Intelligence Age. SSRN Electronic Journal. 96, 87. https://doi.org/10.2139/ssrn.3129557
[18] EU-OSHA (2021). Impact of artificial Intelligence on occupational safety and health. https://osha.europa.eu/en/publications/impact-artificial-intelligence-occupational-safety-and-health
[19] European Commission (2021). Regulatory framework proposal on artificial intelligence. European Union. https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai
[20] Kunst, D. (2020). Deskilling among manufacturing production workers (SSRN Scholarly Paper ID 3429711). *Social Science Research Network.* https://doi.org/10.2139/ssrn.3429711
[21] Surya, L. (2019). Artificial intelligence in public sector. *International Journal of Innovations in Engineering Research and Technology*, *6*(8), 7.

technology or be aware of current regulations, like privacy and data legislation that directly impact AI ventures'[17].

Loss of privacy is another central concern related to the deployment of AI-based systems in workplaces. Widespread data collection is required for AI-based 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 is collected[22]. Failure to develop and enforce ethics guidelines for the collection and use of instruction-related data could result in widespread rights violations.

It must be recognised that **surveillance** can have positive and negative consequences. For instance, in the education sector, monitoring may allow for helpful feedback, student customisation potential, time saving, and so on. But 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[23]. The rising rate of teacher observations as a means to improve education outcomes demonstrates a tolerance and willingness for classroom monitoring, something that AI-based systems could take to whole new levels[24].

Related to a potential loss of privacy is a **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-based systems 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. 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'[25]. In short, restricted choice in the execution of work tasks entails that 'worker autonomy would be undermined, and consequently the jobs' meaningfulness as well'[21].

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. Workers 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. Additionally, this could be associated with a number of health related issues.

Lastly, the literature surveyed for this report, particularly findings from the care and education industries, suggest that the uptake of AI-based systems could induce a process of **depersonalisation**. The introduction of AI-based systems into the care industry is uniquely illustrative. Rubeis explains that the expansion of smart 'technology leads to the distinction between patients as bodies and patients as subjects' because the central focal point of care becomes 'easily measurable indicators that are usually bodily in nature'[26]. In other words, growing involvement of monitoring systems and instructional assistants in the caregiving process transforms the relationship between the carer and patient, ultimately by turning the latter into an object for the former. The patients no longer present their needs as a subject, but instead, their needs are directly observed by the carer through technological devices.

Although the literature tends to focus on the potential benefits of AI-based 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

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

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

[24] 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, https://doi.org/10.1086/700597

[25] 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. https://doi.org/10.1007/s13347-019-00377-4

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

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 in their work. Put another way, the worker no longer extends her[him]self into the decision-making processes, effectively limiting the need to use her emotional and cognitive capacities while providing care. While this can be specifically prevalent in care work, other socially oriented forms of work, like teaching, or even forms of customer service could experience a similar effect.

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 etc., 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.

Conclusion and recommendations

The human impact of AI-based systems are mainly on a psychosocial or organisational level. They should be addressed with the same attention, however, as physical risks.

Any AI-based system in the workplace that collects data should abide by most recent **ethics and privacy** and data protection regulations. In addition, companies that implement AI-based systems should focus on consent, transparency, participation and accountability towards their employees to keep the loss of actual and perceived privacy to a minimum.

Furthermore, active steps need to be taken against deskilling. Not only from the perspective that in case the technology malfunctions or fails, workers might need to perform the task manually, but understanding the work process as part of informed decision-making. It counteracts the feeling of complete dependency on the AI-based system that otherwise could lead to a perceived **loss of autonomy**.

Lastly, when AI-based systems are employed for social tasks, the risks of **depersonalisation** and loss of social interaction among workers and clients, students or patients should not be ignored. Especially in the social field, we see that this technology can automate tasks but not replace the complex layers of human interaction. Countermeasures could provide additional facilitation of social interaction with other people in the workplace.

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).

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

This policy brief 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.