ESCP
BUSINESS SCHOOL

IT ALL STARTS HERE

BERLIN | LONDON | MADRID | PARIS | TURIN | WARSAW

© blacksalmon-AdobeStock

LIGHTS - Human/Well being

Beyond automation: Crafting a future where AI augments, not replaces

ESCP Impact Paper No.2024-45-EN

Louis-David BENYAYER & Hao ZHONG
ESCP Business School

ESCP RESEARCH INSTITUTE OF MANAGEMENT (ERIM)

ESCP Impact Paper No.2024-45-EN

Beyond automation: Crafting a future where AI augments, not replaces

Louis-David Benyayer*
Hao Zhong**

ESCP Business School

Abstract

One major challenge for the years to come, both from a business and a societal perspective, will be to strike a good balance between augmentation and automation of human labor with AI (and Generative AI in particular). On one side, automation has clear limits, like perpetuating biases and lacking adaptability. On the other side, augmentation has clear advantages like improving work quality, job satisfaction, and labor demand. However, both augmentation and automation have negative effects (privacy, security, job destruction, and environmental impact), which should be balanced through laws and regulations. We propose three actions for corporations to navigate the automation/augmentation trade-off and highlight three issues on the European agenda associated with negative externalities triggered by AI technologies.

Keywords: Artificial Intelligence, Automation, Augmentation

*Associate Professor (Teaching), ESCP Business School

** Assistant Professor, ESCP Business School

Beyond automation: Crafting a future where AI augments, not replaces

In the dynamically evolving world of business, the conversation invariably shifts toward the transformative power of artificial intelligence (AI). With rapid advancements shaping the technology landscape, AI has emerged as a disruptive force that reshapes industries and challenges traditional paradigms. Over time, AI technologies play a vital role in driving *automation*, which typically refers to the use of technology to perform tasks and processes without the need for human intervention. It involves the automation of repetitive, rules-based tasks to optimize operations, enhance productivity, and minimize human error using AI algorithms and systems. For example, Klarna, a Swedish fintech company, launched an AI assistant which has shown remarkable effectiveness in handling customer service inquiries. The AI assistant is reported to perform a workload equivalent to that of 700 full-time agents, with projections suggesting a potential $40 million USD in profit improvement in 2024 (Klarna, 2024). However, automation also carries the possibility of adverse effects. Google's newly launched Large Language Model (LLM), Gemini, faced criticism for producing historically inaccurate images that often depicted people of color in situations where this was not appropriate (Robertson, 2024). While automation certainly offers clear benefits, the risks and side effects it poses without human intervention cannot be overlooked.

Alternatively, the concept of *augmentation* emerges as a critical consideration in mitigating the potential pitfalls of automation and ensuring human involvement in the process. *Augmentation* pertains to the use of technology to enhance human capabilities rather than replace them. This strategy prioritizes harnessing AI tools and systems to enhance human decision-making, problem-solving, and even creativity. It seeks to complement human skills and expertise by providing data-driven insights, personalized recommendations, and intelligent assistance to improve overall performance and efficiency. As an example, the development of Sepsis Watch utilized advanced deep learning technologies to enhance the detection and treatment of sepsis, a medical emergency of the human body's extreme and toxic reaction to an infection. In practice, Rapid Response Team (RRT) nurses and Emergency Department (ED) physicians use this tool to support their diagnostic process for sepsis while also exercising their professional discretion. It helps identify patients at risk of sepsis before clinical symptoms manifest, enabling healthcare providers to intervene proactively and improve diagnostic accuracy. The deployment of Sepsis Watch has demonstrated a substantial decrease in sepsis-induced fatalities (Hao 2020).

The debate surrounding the use of automation vs. augmentation has significant implications for corporations. Automation has increasingly become a favored strategy for improving efficiency and productivity in various industries. However, the rise of automation also raises concerns and issues, such as potential job displacement and the erosion of human skills. The argument we develop in this article is that even though augmentation offers a more promising approach than automation for achieving a harmonious balance between AI technology and human intelligence, massive change management efforts are required to reach this objective. While acting at a corporate level is essential, that alone is insufficient to ensure a balance between automation and augmentation that prioritizes the welfare of individuals. From a European perspective, it is crucial for corporations to evaluate the possible negative externalities associated with both approaches to ensure the proper and effective harnessing of AI technologies.

Limits of automation

Let's now delve into a detailed discussion with a focus on the limitations of automation and the potential benefits of augmentation.

Automation perpetuates existing biases and inaccuracies. Automation systems rely on past data and patterns to make decisions. This reliance on historical data means that automation systems can perpetuate existing biases or inaccuracies present in the data they were trained on. Without mechanisms to address and correct these biases, automated systems might further entrench societal inequities or lead to discriminatory outcomes. For instance, if an AI system is used for hiring and is trained on biased historical data with discriminations against certain groups, the system may inadvertently favor applicants from the overrepresented groups.

Automation is less adaptable to unforeseen situations. Automation systems are designed to follow pre-defined algorithms and instructions, which makes them less adaptable to sudden changes or unexpected situations. It may restrict the effectiveness and reliability of automation systems in dynamic environments. This limitation is particularly problematic in fields that require adaptability and nuanced understanding, such as healthcare and disaster response. Conversely, human employees are adept at managing situations that require flexibility, quick decision-making, and adaptability.

Automation overlooks ethical or moral aspects. Automation lacks the ability to incorporate ethical or moral considerations into its decision-making process. Complex decisions that require empathy, ethical judgment, or understanding of human emotions may be beyond the scope of automated systems. This is particularly evident in fields such as the judicial system. Decisions impacting people's lives and freedoms cannot be exclusively delegated to algorithms. They must consider a myriad of factors, including cultural contexts, societal norms, as well as moral responsibility. Such decisions should be approached with a blend of technological aid and human judgment, ensuring that ethical considerations are at the forefront of all endeavors.

These limits of automation are now well identified and some recent developments attempt to address them. For example, there have been initiatives to enhance algorithm development through methods such as adversarial debiasing and causal reasoning to promote AI fairness. Beyond that, the AI Act in Europe stipulates the cases and situations where automated decisions are forbidden.

Benefits of augmentation

When humans are still in the loop, as is the case for augmentation, the above mentioned limits are reduced. Biases stemming from the historical data can be limited by specific measures and criteria for decision making which are driven more by a desired state than by the result of past decisions. Similarly, adaptation to changing conditions is made easier with human intervention. For example, in the context of surgery, prompt human intervention can reduce substantial risks that may arise in the event of automation failure. Last, leaving the final decisions to humans prevents the risk of unethical decisions. An example can be found within the judicial system, where entrusting final decisions to human discretion can alleviate potential injustices arising from pure automation of the trial process.

In addition to reducing the limits of automation, augmentation presents various specific advantages.

Augmentation improves the quality of work. AI augmentation has demonstrated the potential to enhance work quality significantly. This is primarily attributed to the ability of AI systems to identify patterns, trends, and insights that might be overlooked by humans. According to a recent study by the Boston Consulting Group (BCG) (Dell'Acqua et al., 2023), AI-assisted consultants displayed outcomes that were more than 40% superior in quality when compared to a control group lacking AI support. Furthermore, supplementary research underlines the effectiveness of LLMs in offering valuable feedback and generating innovative ideas (Liang et al., 2023), frequently matching or surpassing the capabilities of humans.

Augmentation potentially increases the demand for labor. The innovations of augmentation, which encompass new processes, products, services, or industries, have been found to increase labor demand due to the emergence of a new demand for expert knowledge and specific competencies. These innovations not only enhance the capabilities, quality, variety, or utility of the outputs of occupations but also create a need for a skilled workforce to implement and manage these advancements effectively (Auto et al., 2022).

Augmentation has the potential to enhance job satisfaction. By leveraging AI augmentation, organizations can enhance decision-making processes, leading to increased efficiency and effectiveness. This technology enables employees to dedicate their time to high-level, intricate tasks that require creativity and innovation, fostering a more intellectually stimulating work atmosphere. Augmentation can also support tackling complex problem-solving challenges, thereby alleviating work-related stress and contributing to an overall improvement in job satisfaction. Furthermore, augmentation offers avenues for upskilling and professional growth, empowering employees to engage in continuous learning and development.

Challenges for augmentation

Pure automation falls short of its promises and augmentation has major benefits. However, implementing successful augmentation comes with challenges, three recent research studies illustrate what is at stake.

Rana et al. (2022) surveyed 355 executives about the unintended consequences of using AI for decision-making and their influence on performance. They identify three factors that contribute to a negative effect on performance: poor data quality, lack of governance, and inefficient training. The three factors are grouped under the term “AI opacity,” which has both direct and indirect effects on performance. Direct effects on performance, for example, occur when inadequate training of staff on the systems will result in underutilization of the technology, lower motivation, and a subsequent decline in performance. Indirect effects of this opacity take place when it increases companies’ exposure to risks and suboptimal decisions.

The same phenomenon was described in a recent experimental study by Agarwal et al. (2023). It involved professional radiologists with various availabilities of AI assistance and contextual information to study the effectiveness of human-AI collaboration. The findings reveal that although AI alone was more accurate in diagnoses than two-thirds of radiologists, collaboration with human experts yielded no uniform improvement. The authors explained this by “radiologists' partially underweighting the AI’s information relative to their own and not accounting for the correlation between their own information and AI predictions.” So the authors conclude that “the optimal solution involves assigning cases either to humans or to AI, but rarely to a human assisted by AI.”

A story in the *New York Times* (Jewett, 2023) sheds light on the reasons why some doctors may not follow the recommendations provided by AI systems. In healthcare, AI systems for diagnosis have blossomed and in the U.S., the Food and Drug Administration (FDA) authorized many new programs that use artificial intelligence. However, doctors are skeptical that the tools really improve care or are backed by solid research. “Doctors are raising more questions as they attempt to deploy the roughly 350 software tools that the F.D.A. has cleared to help detect clots, tumors or a hole in the lung. They have found few answers to basic questions: How was the program built? How many people was it tested on? The lack of publicly available information is causing doctors to hang back.”

These three studies illustrate what is at stake for successful augmentation: providing sufficient information about how the AI solution has been built and works, as well as sufficient training on how to interpret the results provided. Augmentation is then as much a technical challenge (designing efficient accurate systems) as a trust and change management challenge (equipping the human working with AI solutions with the information and training required for fostering trust).

Three recommendations for corporations

We can derive from our previous arguments three recommendations for corporations.

Decide strategically on task allocation. A critical aspect of integrating AI into the workforce is discerning which tasks are best suited for automation. The study by Noy & Zhang (2023) illustrates a significant shift in job roles due to the surge of Generative AI, with a notable decrease in time spent on initial drafting and an increase in time devoted to editing. Identifying the right balance between AI and human input becomes crucial for maximizing efficiency and effectiveness. It’s paramount to identify the tasks that can be automated, the ones that should blend AI and human contribution, and the ones which should remain totally managed by humans. Deciding on tasks allocation involves taking into account the technological developments and their cost but also the strategic positioning of the company. Deciding on what AI performs and what is managed by humans is a strategic choice, not a purely technical one. Last, the direct environmental footprint of AI is massive and growing. Electricity and water consumption from data centers are now competing with consumers' needs. The choice of tasks allocation should also be driven by the environmental impact of the AI system used.

Balance investment in technology with investment in human resources. Securing data and infrastructure is not enough when managers are not trained on the AI systems and on how to use the results they produce. Training is becoming increasingly vital, especially to enhance the skills of lower performers. As AI continues to commoditize certain aspects of knowledge work, it’s not just the low performers who are affected; skilled professionals are also at risk. Comprehensive training programs can help employees understand the nuances of AI, including its limitations and best practices (Brink et al., 2023). The synergy between technical and human assets is key to realizing the full potential of AI. Success in this domain not only depends on redefining processes but also hinges on fostering an organizational culture that understands and trusts AI.

Engage with stakeholders in a conversation about the use and conditions of automation and augmentation. Significant efforts should be put into ensuring an understanding of the system design by the users. Transparency on the data used, the models implemented, the results' limits and the scope of relevance all foster adoption through higher transparency. What also supports adoption is sharing a common view about the tasks allocation we described earlier and developing a common sense of the split between automation and

augmentation. More precisely, these discussions could aim to decide on the details of augmentation, for example, to state the right level of transparency according to the context (e.g., facial recognition for smartphone vs radiology) or the objective (making the decision vs performing a task). These conversations should also involve external stakeholders such as clients and partners. Unions and collectives should also be involved to determine the scope and modalities of automation and augmentation as was the case in the U.S. in September 2023 when an agreement was reached between the show writers and the studios about the use of AI in the writing process.

Three issues on the European agenda to balance negative externalities

Acting at a corporate level is important and useful but is not enough to ensure a balance between automation and augmentation that favors citizens. This is because automation and augmentation with AI generate negative externalities for which corporations are not naturally held accountable for. Whether it's bias, privacy, security risks, misinformation, job destruction, market concentration, or environmental impact, the impacts of AI on society go far beyond the corporate world and law. Regulations are a way to reduce and manage these negative externalities.

The program of the future commission is not yet known but we can already mention 3 points of attention which we encourage the future commission, parliament, and civil society at large to take into consideration in the future discussion.

Protect European citizens against AI-related risks. Privacy, bias and discrimination, cybersecurity, and misinformation are identified risks (Sartor, 2019) that are in the scope of voted regulations (Digital Market Act, Digital Services Act, AI Act, Cybersecurity Act, Data Governance Act). Monitoring regulations application and impact will be critical, as well as identifying new regulations to put on the agenda to adapt to new technological developments.

Prepare the European workforce and ensure jobs quality and quantity in Europe. AI has a strong impact on the job market. It has the power to reduce demand for some jobs, increase it for others and transform a lot of them. This calls for a massive plan to prepare the European workforce: identifying jobs at risk and planning for upskilling the concerned workers; building expertise in AI through investment in research and teaching; training the workers massively about the technology, its use and limits; implementing rules and regulations which will preserve job quality in Europe. On this last point a major challenge lies in the global reduction in job quality triggered by AI: good-quality jobs are replaced by massive needs in data labeling and lower analysis.

Reduce AI environmental footprint. Training large AI models requires significant computational power, contributing to energy consumption and carbon emissions. The environmental footprint of developing and running AI systems is a growing concern, particularly as these technologies scale. Reducing environmental footprint would entail supporting investments in IT systems efficiency but more importantly developing policies and actions that would consider and actively mitigate rebound and induction effects. This means integrating AI solutions with a clear vision of promoting sustainable consumption patterns.

Conclusion

As the technology and the usage develop, it becomes increasingly clear that the path forward is not one of automation or augmentation in isolation, but a judicious blend of both. This journey demands a conscious effort to balance technological advancements with ethical considerations, inclusivity, and environmental sustainability. Strategic decision-making lies at the heart of this balance. The future of AI in business and society hinges on a collaborative dialogue among all stakeholders—policymakers, corporations, and the broader public – to navigate the challenges posed by AI, from ethical dilemmas to socio-economic impacts. The journey ahead is complex and requires careful stewardship, AI is not only a technological matter but a governance one.

References

Agarwal, N., Moehring, A., Rajpurkar, P., & Salz, T. (2023). Combining Human Expertise with Artificial Intelligence : Experimental Evidence from Radiology (Working Paper 31422). *National Bureau of Economic Research*. https://doi.org/10.3386/w31422

Autor, D., Chin, C., Salomons, A. M., & Seegmiller, B. (2022). New frontiers: The origins and content of new work, 1940–2018 (No. w30389). *National Bureau of Economic Research*.

Brink, A., Benyayer, L.-D., & Kupp, M. (2023). Decision-making in organizations : Should managers use AI? *Journal of Business Strategy, Vol. 45 No. 4* . https://doi.org/10.1108/JBS-04-2023-0068

Dell'Acqua, F., McFowland, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., ... & Lakhani, K. R. (2023). Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality. *Harvard Business School Technology & Operations Mgt. Unit Working Paper*, (24-013).

Hao, K. (2020). How an AI tool for fighting hospital deaths actually worked in the real world. *MIT Technology Review*. Retrieved April 6, 2024, from https://www.technologyreview.com/2020/10/02/1009267/ai-reduced-hospital-deaths-in-the-real-world/

Jewett, C. (2023, octobre 30). Doctors Wrestle With A.I. in Patient Care, Citing Lax Oversight. *The New York Times*. https://www.nytimes.com/2023/10/30/health/doctors-ai-technology-health-care.html

Klarna (2024). Klarna AI assistant handles two-thirds of customer service chats in its first month | Klarna International. Retrieved April 6, 2024, from https://www.klarna.com/international/press/klarna-ai-assistant-handles-two-thirds-of-customer-service-chats-in-its-first-month/

Liang, W., Zhang, Y., Cao, H., Wang, B., Ding, D., Yang, X., ... & Zou, J. (2023). Can large language models provide useful feedback on research papers. A large-scale empirical analysis. In arXiv preprint.

Noy, S., & Zhang, W. (2023). Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence (SSRN Scholarly Paper 4375283). https://doi.org/10.2139/ssrn.4375283

Rana, N. P., Chatterjee, S., Dwivedi, Y. K., & Akter, S. (2022). Understanding dark side of artificial intelligence (AI) integrated business analytics : Assessing firm's operational inefficiency and competitiveness. *European Journal of Information Systems*, 31(3), 364387. https://doi.org/10.1080/0960085X.2021.1955628

Robertson, A. (2024). Google apologizes for "missing the mark" after Gemini generated racially diverse Nazis. *The Verge*. https://www.theverge.com/2024/2/21/24079371/google-ai-gemini-generative-inaccurate-historical

Sartor, G. (2019). Artificial Intelligence : Challenges for EU citizens and consumers. European Parliament.