

THE HUMAN FACTOR IN ARTIFICIAL INTELLIGENCE

CHALLENGES

AI biases. During data collection or algorithm design, humans may impart their preferences, prejudices and cognitive biases, resulting in AI systems that perpetuate or exacerbate existing inequalities.

Data Opacity. Complex AI models make decisions that don't follow the rules set by humans, through hard-to-decipher algorithms, which can affect the lives of people and communities.

RISKS

AI could negatively impact 35% of the Sustainable Development Goals and exacerbate inequalities.

It could lead to risks for 30% of jobs in Portugal and create information monopolies (and thus lower quality and respect for privacy, as well as behavioural manipulation).

It could lead to other problems, such as AI-based weapons and misalignment of goals between AI and humans.

BENEFITS

AI could positively impact 79% of the Sustainable Development Goals.

It could be beneficial to people as regards potentially improving access to healthcare and education.

It could have benefits for society, such as creating 97 million new jobs and increasing productivity by between 11% and 37%.

Artificial Intelligence (AI) has a transformative potential across all sectors of society, but it needs to be regulated to overcome its challenges, leverage its benefits and mitigate its risks.

Regulation should be informed by multiple areas of knowledge, including Psychological Science, to promote development in line with the Sustainable Development Goals (SDGs).

STRATEGIC RECOMMENDATIONS FOR SUSTAINABILITY



Safety, Health and Well-being. Consider different variables, including psychological and social ones, in the development of algorithms that govern the functioning of AI systems.



An Ethic of Transparency. Informing and clarifying – making visible – the inputs and parameters of the algorithms used in AI systems is an ethical responsibility.



Equity, Diversity and Inclusion. Ensure that AI can benefit everyone, with no group being excluded or disadvantaged. AI should be trained for diversity and inclusion, by diverse and inclusive teams.



Human Involvement. Investigate and monitor the interaction between humans and AI systems, implemented in different contexts, prioritising the development of human-centred AI systems, and seeking autonomous and intelligent technologies to complement the work of professionals, improving productivity and the quality of services.



Privacy and Regulation. Promote regulation by helping users make informed decisions regarding the sharing of their personal information, through informed consent, taking into account confidentiality, anonymity, utility and cybersecurity with regard to data.



Literacy Training and Promotion. A commitment to training professionals from different fields, as well as all citizens, is essential to ensure that people will use and apply AI systems productively, responsibly and ethically, maximising their benefits and mitigating their risks.



Contributions from Psychological Science. Codify the contributions of Psychological Science in AI systems and models from their initial design through to implementation.



Financiamento. Priorizar o financiamento de investigações que articulem o contributo da Psicologia no desenvolvimento e aplicação de sistemas de IA.

