Why we should trust AI?

José Domínguez[1], Ayishwarya Laksumi[2], and Sumaiya Sultana[3]

[1] Université Toulouse Capitole 1, Toulouse 31000, France
jose.dominguez@ut-capitole.fr
[2] Université Toulouse Capitole 1, Toulouse 31000, France
ayishwaya-laksumi.ganeshraam@ut-capitole.fr
[3] Université Toulouse Capitole 1, Toulouse 31000, France
sumaiya.sultana@ut-capitole.fr

Abstract. In modern society, Artificial Intelligence (AI), is becoming increasingly predominant in decision-making for various domains such as healthcare or finance. Nevertheless, trust in AI-based systems remains a significant issue. This paper explores the reasons why we should trust in AI and possibly achieve its successful implementation. By examining the factors contributing to trust, including transparency, reliability, and accountability, we argue that fostering trust in AI is crucial for accepting the arrival of a new technological revolution. Through an exhaustive analysis, we highlight the benefits of trusting AI and propose strategies to lessen its flaws, ultimately paving the way for complete integration into today's world.

Keywords: AI · Trust · Artificial Intelligence · Reliability

1 Introduction

Artificial Intelligence is changing the global landscape, emerging as a pivotal force in the digital domain and driving innovation across various sectors. For instance, Amazon (US) has been using AI for predictive logistics for several years now, having patented the technology back in 2014. The online retail giant analyses customer data to predict demand for goods[7]. Not only big companies use it as a way to automatize routine processes but also conventional users.

It is becoming very normal the use AI chatbots like ChatGPT to ask about general information, grammar corrections, homework doubts, or even recipes. A

Forbes Advisor poll found that 85% of Britsh are aware of AI language models such as Chat GPT, Google Bard, and Bing Chat [6].

So, in this rapidly growing technological era, our debate is whether we should trust AI (Artificial Intelligence) or not. There is indeed a lot of confusion or uncertainties regarding trusting AI. In this paper, we will explore why people should trust AI not only for being reasonable but also because it is very urgent for progress and innovation.

Trusting AI is related to many critical questions, such as: what are the main barriers to trusting AI? On what basis can we trust AI? This debate is very important as it is closely related to the reliability, safety, and ethical implications of integrating AI, which affect our daily lives in various aspects such as healthcare, finance, transportation, etc. On the other hand, distrust can lead to a situation where it might hinder the progress of innovation, and people in society might miss opportunities.

Our analysis in this debate will focus on a comprehensive understanding of AI, its capabilities, limitations, and ethical considerations. It should also consider the expertise of researchers, scientists, and professionals working in the field of AI, who are continually striving to enhance its safety, reliability, and ethical framework.

2 Exposition

In this debate, our position is in favor of trusting AI. We believe that with proper safeguards and ethical guidelines, AI has the potential to revolutionize industries, solve complex problems, and improve the human condition. We will prove it with countless examples of AI being used beneficially in various sectors.

3 Arguments

3.1 Algorithmic Biases

Computers need data to learn from. Examples of training data are text, images, and computer code. In most cases, the larger the data set, the better the AI will perform. But no data set is perfectly objective; each comes with potential biases or preferences. Not all biases are unjust, but the term is most often used to indicate an unfair advantage or disadvantage for a certain group of people.

Data Analysis: a possible approach to solve this issue is to eliminate the potential for bias before any AI system is deployed. This type of task could be done by independent entities rather than companies. They normally have internal interests and usually the urge to deploy their technology as soon as possible. Distributed Artificial Intelligence Research Institute publishes best practices that could be adopted by industry. For instance, they propose including a data-sheet in the training datasets. It would include "its motivation, collection process, recommendations.."[8]

Fairness-Aware Machine Learning Techniques: There exist techniques where it can ensure fairness in machine learning techniques, for example, the one proposed by Zliobaite (2015). It can identify the bias in AI systems, promoting fairness and equity.

3.2 Building Trustworthy AI Through Transparency and Fairness

While claiming the immense potential of AI, some people still show hesitation in trusting it due to lack of transparency, reliability, and bias concerns. Instances of AI algorithms exhibiting racial or gender bias highlight the importance of fair and equal treatment in AI development. Additionally, the lack of transparency in complex AI systems, often referred to as "black boxes," makes it difficult to understand how they reach decisions, hindering trust.

However, these concerns are not roadblocks to progress. Researchers are actively developing methods to improve the transparency of AI models, allowing us to gain a clearer picture of their decision-making processes. This will significantly enhance trust in AI. Furthermore, efforts are on the way to mitigate bias by employing diverse training data sets that reflect the richness of populations.

" You have to treat all AI like a community, a society," says Mory Gharib, a researcher in Caltech University. He states the necessity of protocols, like laws in our society, that future AI technologies can not violate to make sure that these systems can not affect negatively us or a third party.

As technology continues to evolve, these safeguards will become increasingly robust. This will foster a more trustworthy environment where humans and AI can collaborate effectively, paving the way for a future enriched by responsible and beneficial integration.

3.3 Job Loss Problem

The job loss problem due to AI automation is one of the main concerns in distrusting AI. The figures about job loss presented from various studies, such as the Oxford Study and report from the World Economic Forum and McKinsey, may represent worst-case scenarios or speculative projections (Frey & Osborne, 2017). So, it is highly important to approach those figures in the report presented with caution and consider alternative perspectives.

Dynamic Nature of Job Markets: AI can lead to certain jobs disappearing. However, it also opens up many new job opportunities. Looking back in history, whenever technology advances, it tends to create new jobs too, making up for the ones that were lost (Acemoglu & Restrepo, 2019). This means that the impact of AI on jobs might not be as simple as some people fear.

Potential for Skill Enhancement and Adaptation: Improving the education system and emphasizing skills like thinking critically, being creative, and innovating can help people adapt to changes in the job market (Brynjolfsson & McAfee, 2014). Instead of seeing AI as something to worry about, we can focus

on teaching new skills and improving existing ones. Ultimately, this way, people can stay competitive as the job market changes. So, investing in developing people's skills can also help them switch to different types of jobs and industries.

Opportunities in the Gig Economy(free market system): The gig economy offers chances for people to earn extra money or start new careers (Katz & Krueger, 2016). For instance, websites and apps that offer freelance work allow people to use their skills in flexible ways. Indeed, joining the gig economy can help people feel more financially secure and better able to handle changes caused by AI technology.

4 Conclusion

In conclusion, trusting AI is crucial for embracing technological advancements and fostering innovation. While concerns like algorithmic biases and potential job losses exist, they can be addressed through transparency, fairness-aware machine learning techniques, and skill enhancement programs. By implementing safeguards and ethical guidelines, we can ensure the responsible integration of AI into various domains.

Moreover, recognizing the dynamic nature of job markets and leveraging opportunities in the gig economy can help individuals adapt to unexpected changes. Embracing trust in AI opens doors to countless possibilities, ultimately leading to a future technology era enriched by responsible and beneficial AI integration.

References

1. Acemoglu, D., & Restrepo, P. (2019). Automation and New Tasks: How Technology Displaces and Reinstates Labor. NBER Working Paper No. 25684.
2. Autor, D. H. (2015). Why Are There Still So Many Jobs? The History and Future of Workplace Automation. *Journal of Economic Perspectives, 29*(3), 3-30.
3. Brynjolfsson, E., & McAfee, A. (2014). The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies. W. W. Norton & Company.
4. Frey, C. B., & Osborne, M. A. (2017). The Future of Employment: How Susceptible Are Jobs to Computerisation? *Technological Forecasting and Social Change, 114*, 254-280.
5. Katz, L. F., & Krueger, A. B. (2016). The Rise and Nature of Alternative Work Arrangements in the United States, 1995-2015. National Bureau of Economic Research.
6. Forbes. (2024, April 12). UK Artificial Intelligence (AI) Statistics And Trends In 2024. Retrieved from https://www.forbes.com/uk/advisor/business/software/uk-artificial-intelligence-ai-statistics-2023/
7. Economist Intelligence. (2023, June 5). How companies use artificial intelligence. Retrieved from https://www.eiu.com/n/how-companies-use-artificial-intelligence/
8. Caltech. (n.d.). Can we trust artificial intelligence? Retrieved from https://scienceexchange.caltech.edu/topics/artificial-intelligence-research/