

Artificial intelligence risks to humans

Student name/ Rakan Fahad Aloufi

University ID number/ 2021205012

Graduation Project – 1

Faculty of Management - Midocean University

Supervisor

Dr. Mona Moustafa Elashry

Faculty member- Faculty of management- Midocean University

1 st Semester- 2024

Table of contents

Chapter 1:

Introduction:	3
<i>Research Problem:</i>	4
The theoretical gap 2.1:	4
Empirical gaps 2.2:	4
<i>Research Significant:</i>	5
The theoretical significant 3.1:	5
Empirical significant 3.2:	5
Research Sub-Questions:	6
Research Objectives:	7
Research Hypotheses:	8
Research Model:	10
Empirical Study:	11
Research Methodology :	13
References list :	14

Introduction

During the past five decades, the dangers that many people were warning about were population explosion, the decrease in natural resources and minerals, environmental pollution, and the possibility of a nuclear war. However, today, with the tremendous development in the level of technology and its applications, the matter has turned to other dangers. Robots whose intelligence will reach a degree that allows them to outperform humans, especially with artificial intelligence capabilities, or devastating computer viruses that stop the Internet from working. Artificial intelligence is the new fad that is being talked about in all fields.

A research paper issued by experts in artificial intelligence in industrial and academic fields, entitled “Harmful Use of Artificial Intelligence: Prediction, Prevention, and Mitigation,” warned that this sector will negatively affect the human race and will at some point lead to serious damage to several sensitive aspects of its life. The study considered that “dual use” and the ability of technical devices to make thousands of complex decisions per second can be used in both extremes, either to benefit or harm people, and this depends on the person who designs the system. The experts took into account the uses of artificial intelligence that currently exist or that could be developed during the next five years, and were then divided into 3 groups: digital, physical and political.

Machines equipped with special capabilities will reach levels of intelligence that exceed human levels, enabling them to make decisions independently, without referring to any human reference. Google recently appointed an ethics council to deal with deceptive issues related to artificial intelligence and help the company confront criticism related to its deals in this field. However, in less than a week, this council began to disintegrate amid resignations from a number of its members, which could pose a threat to Google's chances of winning more military cloud computing contracts.

A number of council members, including behavioral economics expert Alessandro Acquisti, announced that he would not continue in his duties assigned to him by Google. Today, a number of employees at Google began submitting a petition to the company to oust more members of this council, amid strong opposition to the technology company's involvement in military contracts to develop weapons with artificial intelligence.

The risks of AI are great, but the use and growth of these tools is also inevitable, as the benefits go beyond simple efficiency gains and include a fairer decision-making scenario when algorithms are trained to avoid bias, while also increasing our understanding from a risk management and auditing perspective, where we should look for the key to advantage. In artificial intelligence systems. Control and governance of AI must take precedence over algorithms and efficiency. We all have a responsibility to learn more about AI risks and control those risks. The topic is not going away. Risks will continue to grow and change as the technology becomes more advanced and widespread.

Research problem

If something has a bright side, then in parallel, it also has a dark side. Likewise, artificial intelligence also has some drawbacks. Although AI is one of the most popular and demanding technologies around the world, it still has some drawbacks and problems.

The theoretical gap 2.1:

You need to produce new AI-powered software at a financially significant cost, as well as update existing sports software to keep up with the latest requirements, in addition to the cost of repair and maintenance. Robots are one of the applications of artificial intelligence, and they have begun to replace humans in some jobs. Likewise, for artificial intelligence applications for various writing, image editing, etc., starting from jobs, it should not appear in those digital jobs. Humans have begun to rely on artificial intelligence to perform many tasks, especially repetitive ones. They have also begun to rely on robots to complete some work, and it is expected that humans will be completely dependent on artificial intelligence in the future. It is one of the risks of artificial intelligence. Advanced AI technology needs a large amount of data for training and analysis, and this raises concerns about data privacy and its proper use. Personal data may be exploited in illegal or unethical ways, putting the rights of individuals and communities at risk.

Empirical gaps 2.2:

Although artificial intelligence has come a long way in the world of medicine, human monitoring is still necessary. For example, surgical robots act logically, rather than empathically. Health practitioners may observe biobehavioral observations, which can help diagnose or prevent medical complications. The integration of artificial intelligence technology is growing rapidly in education, due to its great benefits, as it improves student participation and interaction by enhancing the learning experience in different ways and directing teachers to use a more interactive educational approach. Artificial intelligence tools can also participate in evaluating students and analyzing their data to become Things are clearer in front of the teacher.

Main research question

This study addresses the following main research question:

“How does artificial intelligence affect human lives?”

Research Significant

AI is used to improve efficiency and productivity in many fields, such as manufacturing, healthcare, and marketing. For example, AI can be used to improve the accuracy and speed of manufacturing processes, to improve medical diagnosis or develop new treatments, or to improve advertising targeting.

The theoretical significant 3.1:

The recipe for artificial intelligence is to help save time and effort gradually, as the exercise can be performed faster and faster than a human. For example, AI can be used to prepare reports, analyze data, or provide assistance with household tasks.

Empirical significant 3.2:

Artificial intelligence can help improve human quality of life in many aspects, such as health care, education, and entertainment. For example, AI can be used to provide better health care, develop new ways to educate children, or create new entertainment experiences. Artificial Intelligence is a powerful technology that has the potential to significantly impact human life in many aspects. As artificial intelligence continues to develop, the importance and influence of this technology is expected to increase in the coming years.

Research Sub-Questions

What is artificial intelligence?

What are the risks of artificial intelligence on humans?

What are risks artificial intelligence in Real-life ?

Why research artificial intelligence safety?

Research Objectives

artificial intelligence is a specific branch of computer science concerned with mimicking human thinking and decision-making processes. These programs can often revise their own algorithms by analyzing data sets and improving their own performance without needing the help of a human. These are often programmed to complete tasks that are too complex for non-AI machines.

We talked briefly about real-life and hypothetical AI risks above. Below, we've outlined each in detail. Real-life risks include things like consumer privacy, legal issues, AI bias, and more. And the hypothetical future issues include things like AI programmed for harm, or AI developing destructive behaviors.

There are a myriad of risks to do with AI that we deal with in our lives today. Not every AI risk is as big and worrisome as killer robots or sentient AI. Some of the biggest risks today include things like consumer privacy, biased programming, danger to humans, and unclear legal regulation.

Not that many years ago, the idea of superhuman AI seemed fanciful. But with recent developments in the field of AI, researchers now believe it may happen within the next few decades, though they don't know exactly when. With these rapid advancements, it becomes even more important that the safety and regulation of AI be researched and discussed at the national and international levels.

In 2015, many leading technology experts (including Stephen Hawking, Elon Musk, and Steve Wozniak) signed an open letter on AI that called for research on the societal impacts of AI. Some of the concerns raised in the letter cover things like the ethics of autonomous weapons being used in war, and safety concerns around autonomous vehicles. In the longer term, the letter posits that unless care is taken, humans can easily lose control of AI and its goals and methods.

The importance of AI safety is to keep humans safe and to ensure that proper regulations are in place to ensure that AI acts as it should. These issues may not seem immediate, but addressing them now can prevent much worse outcomes in the future.

Research Hypotheses

Main hypothesis H1:

Artificial intelligence (AI) refers to computer systems capable of performing complex tasks that historically only a human could do, such as reasoning, making decisions, or solving problems

Today, the term “AI” describes a wide range of technologies that power many of the services and goods we use every day – from apps that recommend tv shows to chatbots that provide customer support in real time. But do all of these really constitute artificial intelligence as most of us envision it? And if not, then why do In this article, you’ll learn more about artificial intelligence, what it actually does, ‘we use the term so often and different types of it. In the end, you’ll also learn about some of its benefits and dangers and explore flexible courses that can help you expand your knowledge of AI even further.

The first sub-hypothesis Ha:

In my view, Artificial Intelligence (AI) will be the most disruptive technology over the next decade. The quality of this technology has improved considerably in a number of areas in recent years, with all the consequences this entails. Smart software systems gain an ever better understanding of who we are, what we do, what we want and why we want it. A world full of opportunities opens up. Chatbots, smart virtual assistants and autonomous intelligence software assistants will increasingly come to our aid. AI systems will bring us prosperity, time savings, convenience, insights and comfort. We will become used to a personal assistant that is available 24 hours a day and knows what we need, before we know it ourselves. Just as it is difficult to imagine living without the Internet, a decade from now will see the same scenario in terms of your personal assistant.

Also, smart AI systems will provide us with insights that we believed would never be possible, and they will provide answers to questions the existence of which we were not aware of. AI systems are faster, are never tired, learn from examples and from each other and are considerably smarter than humans in specific domains. This is not a futuristic idea but reality.

The second sub-hypothesis Hb:

The so-called human-in-the-loop approach is especially crucial in sectors where AI-based decisions directly impact individual lives and society .These sectors include:

Education: As demonstrated by the UK, AI systems should not be solely responsible for grading assignments or predicting students’ academic performance. A human teacher’s expertise and personal understanding of their students should play a decisive role in these situations.

Healthcare: AI has made significant strides in disease diagnosis, treatment planning, and patient care. However, the potential for misdiagnosis or inadequate treatment planning due to biases or errors in AI .systems emphasizes the necessity of human professionals in the final decision-making process

Recruitment and HR: AI is increasingly being used for resume screening and predicting potential job performance. However, reliance on AI can lead to biased hiring practices, potentially overlooking candidates

with unconventional backgrounds or skill sets. A human-in-the-loop approach ensures comprehensive and fair candidate evaluation.

Finance and Lending: AI algorithms can evaluate creditworthiness, but they may inadvertently discriminate based on geographical location or personal spending habits, which can correlate with ethnicity or socioeconomic status. In such scenarios, human judgment is necessary to ensure balanced lending decisions.

Criminal Justice: AI is being used to predict crime hotspots and potential reoffending. However, bias in historical crime data can lead to unjust profiling and sentencing. Human oversight can provide more nuanced perspectives and help prevent such injustices.

Autonomous Vehicles: Though AI drives the operation of self-driving cars, it's crucial to have a human in the decision-making process, especially when the vehicle has to make ethical decisions in scenarios of unavoidable accidents.

The third sub-hypothesis Hc:

The understanding, prevention, and mitigation of harms from AI. These harms could be deliberate or accidental; caused to individuals, groups, organisations, nations or globally; and of many types, including but not limited to physical, psychological, social, or economic harms.

Artificial intelligence

Our lives:

- Schools
- the health
- industry
- commerce

Positives:

- Error-free processing
- Helps with repetitive jobs
- Always available
- Make the right decisions
- Speed of making decisions

Negative:

- High costs
- Increased unemployment
- Poor creativity
- Security and privacy risks
- Increased laziness among individuals

Empirical Study

Introduction

In 1956, at a conference at Dartmouth University, scholars formally proposed the term "artificial intelligence." That moment was the first step in a new topic of studying how machines simulate human intelligent activities. In early 2016, AlphaGo defeated the world chess champion. This event immediately aroused global interest in artificial intelligence (AI) [93]. The development of artificial intelligence has brought huge economic benefits to mankind and has benefited all aspects of life, even as it has greatly promoted social development and brought social development into a new era [57]. Many scholars started AI-relevant research since the end of 20th century [92].

AI is the general term for the science of artificial intelligence. It uses computers to simulate human intelligent behaviors and it trains computers to learn human behaviors such as learning, judgment, and decision-making [94]. AI is a knowledge project that takes knowledge as the object, acquires knowledge, analyzes and studies the expression methods of knowledge, and employs these approaches to achieve the effect of simulating human intellectual activities [19]. AI is a compilation of computer science, logic, biology, psychology, philosophy, and many other disciplines, and it has achieved remarkable results in applications such as speech recognition, image processing, natural language processing, the proving of automatic theorems, and intelligent robots [20]. AI plays an indispensable role in social development, and it has brought revolutionary results in improving labor efficiency, reducing labor costs, optimizing the structure of human resources, and creating new job demands [18].

The rest of the paper is structured as follows (Fig. 1): the origin and development of AI is addressed in Section II. Section III describes enabling drivers and technologies of AI. The research fields are depicted in Section IV. Section V illustrates the application scenarios of AI. Three logical views of AI are explained in Section VI, and Section VII summarizes the paper.

Section snippets

The origin and development of artificial intelligence

Artificial intelligence is the study of how to make computers perform intelligent tasks that, in the past, could only be performed by humans [36]. In recent years, AI has developed rapidly, and it has changed people's lifestyles [37]. The development of AI has become an important development strategy for countries around the world, enhancing national competitiveness and maintaining security [69]. Many countries have introduced preferential policies and have strengthened the deployment of key

Big data

Big data is a prerequisite for AI, and it is a core factor that promotes AI to improve recognition rate and accuracy. With the development and the wide application of the Internet of Things, the amount of data generated has increased exponentially, with a great increased annual growth rate. In addition to increasing the number, the dimensionality of the data has also been expanded [105]. These large amounts of high-dimensional data make the data more comprehensive and more sufficient to support

The expert system

An expert system is a knowledge system based on the existing knowledge of human experts. The expert system was the earliest field AI research. It is widely used in medical diagnosis, in geological surveying, and in the petrochemical industry. Expert systems usually refer to various knowledge systems [60]. This is an intelligent computer program, based on knowledge, that uses professional knowledge provided by human experts to simulate the thinking process of human experts and uses knowledge and

Application scenarios of ai

Based on the relatively mature development of technical conditions such as data, algorithms, and computing capabilities, AI has begun to truly solve problems and to effectively create economic benefits [116]. From an application perspective, industries with a good data foundation (such as finance, healthcare, automotive, and retail) have relatively mature AI application scenarios.

Three viewpoints of AI

Symbolism, connectionism, and behaviorism represent the three main viewpoints of concepts in the field of AI research. They are the most important theoretical theorems in the development of AI disciplines, and they are the theoretical foundation for the development of AI disciplines.

Conclusion

The study presents a systematic overview of AI that focuses on prospects and development, core techniques, applicable scenarios, and challenges. This paper conducts a state-of-the-art review of the extant and upcoming research on AI.

Artificial intelligence is an interdisciplinary subject that involves information, logic, cognition, thinking, systems, and biology. It has been used for knowledge processing, pattern recognition, machine learning, and natural language processing. Applications have

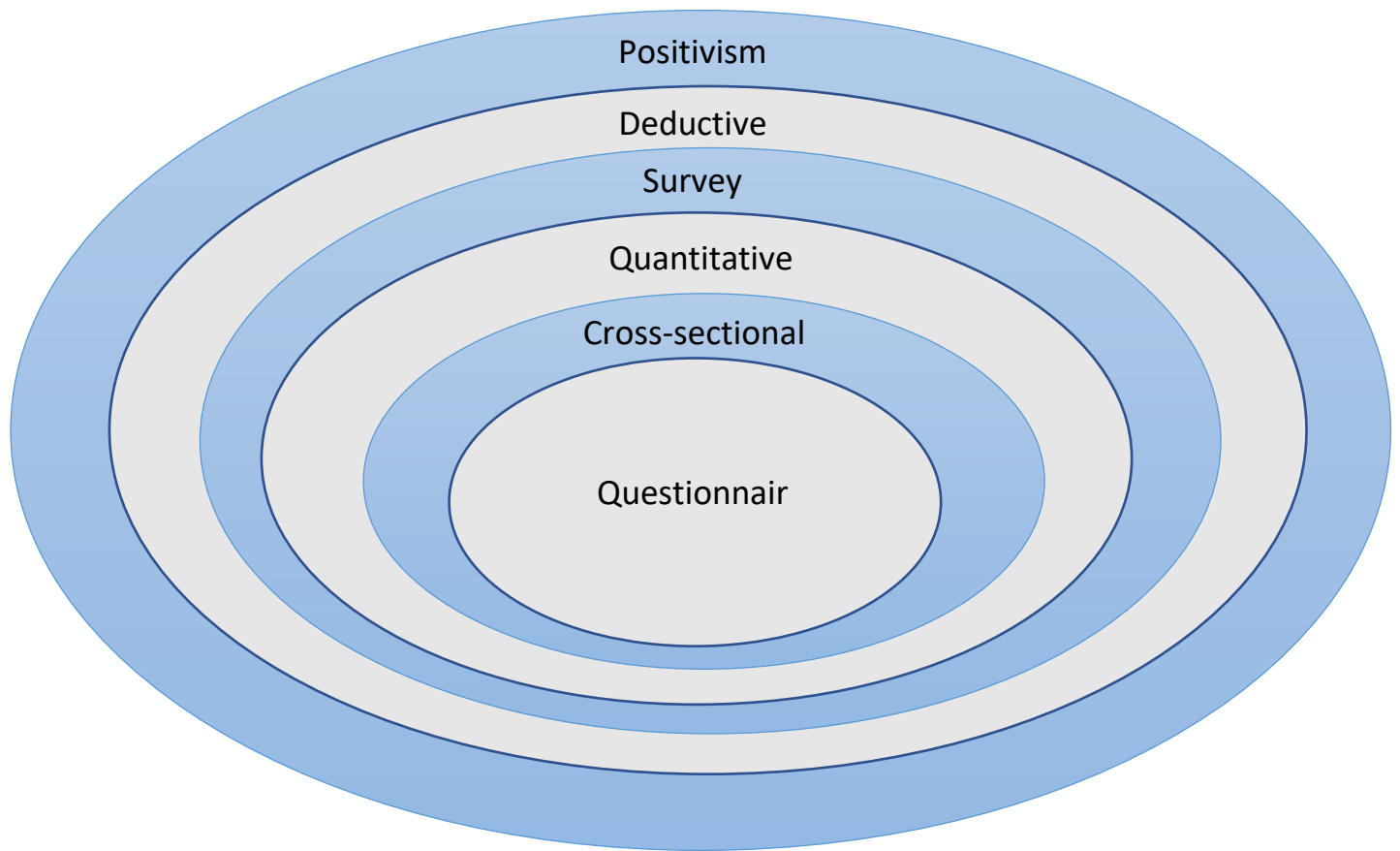
Declaration of Competing Interest

We declare that we have no financial and personal relationships with other people or organizations that can inappropriately influence our work, there is no professional or other personal interest of any nature or kind in any product, service and company that could be construed as influencing the position presented in, or the review of the manuscript entitled A review of Artificial Intelligence: The State of the Art and Future Prospects.

Acknowledgment

Funding of China National Social Science (No. 15BJY022), Funding of Researching of China University of Labor Relations (No. JG2106, No. ZYTS201806)

Research Methodology



References list

Books:

1. Science Fiction and Philosophy: From Time Travel to Superintelligence.
2. Artificial Intelligence: A Guide for Thinking Humans

Magazines:

1. Adams, S., Arel, I., Bach, J., Coop, R., Furlan, R., Goertzel, B., & Sowa, J. F. (2012). Mapping the landscape of human-level artificial general intelligence. *AI Magazine*, 33, 25–42.
2. Goertzel, B. (2014a). GOLEM: towards an AGI meta-architecture enabling both goal preservation and radical self-improvement. *Journal of Experimental & Theoretical Artificial Intelligence*, 26, 391–403.
3. Kornai, A. (2014). Bounding the impact of AGI. *Journal of Experimental & Theoretical Artificial Intelligence*, 26, 417–438.
4. Yudkowsky, E. (2008). Artificial intelligence as a positive and negative factor in global risk. *Global catastrophic risks*, 1(303), 184.