International Journal of Research and Publication
ISSUE 5 VOL. 3 (MAY 2024)

Navigating the AI Landscape: History, Types, Terminologies, Ethical Considerations, and Future Impact

[1]*G.O. Asoronye, [2]E. Okekenwa, [3]J. E. Okorie, & [4]Nwaguiyi Jeremiah Theophilus
[1]Computer Engineering Department
[2&3]Electrical Electronic Engineering Department
[4]Computer Science Department
[1&2]Akanu Ibiam Federal Polytechnic Unwana, Ebonyi State, Nigeria
[3]Federal Polytechnic Oko, Anambra State, Nigeria
Corresponding email: *asorgay@gmail.com
Corresponding phone contact: *+2348069145552

Abstract

This paper contemplates the potential changes in the world by using artificial intelligence (AI), it clarifies its history and its types, names and its ethical consideration and choices in the future. Starting with its origin from symbolic reasoning to modern machine learning and deep learning, differentiating Narrow AI (such as natural language processing), General AI systems that seek to replace all human cognitive abilities, and Superintelligent AI which surpasses human intelligence in infinite possible domains are interpreted. For instance, there is deep learning applied to producing huge historical timelines by interpreting ancient damaged inscriptions in the past. Ethical consideration is urgently needed as there are cases of algorithmic bias influencing the understanding in casting history. And, meanwhile, the conclusion answers the final question: the use of AI in working histories still has its question and challenge by technologists and data scientists, while it is a promising development in history-making which may help to bridge the gap between worldwide global historians. The lessons that AI has thought us today could be equally distorted at the same time. On the one hand, the answer is uncertain: whether the cultural heritage could be preserved in time; on the other hand, the achievement is: digitalization is a wonderful gap-bridging and AI can help us to understand the world. Over clear and comprehensive explanation for average reader, this paper supplies a digestible framework of vision in AI of history, from the past to the future, which really pays attention to balancing its opportunities and challenges, or making critical reflection about its continuous development.

Keywords: Artificial Intelligence, Machine Learning, Ethics in AI, Superintelligent AI, Natural Language Processing.

Introduction

Artificial Intelligence (AI) has become a disruptive force in just about every sector, fundamentally changing methodologies and outcomes in almost everything we do. Most fields – from healthcare and finance to education and retail – have adapted forms of AI that accelerate time to value by improving efficiency, accuracy and decision-making. A brief history of AI shows that the field began in the mid-20th century with attempts at symbolic reasoning and problem solving, which, however, could only be carried out by humans (Ogbonnia, 2017). As technology progressed, it advanced to machine learning, and later to deep-learning schemes that process data to carry out specific tasks, including personalized medicine and predictive analytics. (Topol, 2019; Esteva et al., 2019). Another all-encompassing example is healthcare, where AI is widely applied for diagnostics, prognostics, patient management (Jiang et al, 2017), medical image analysis for detection of diseases such as cancer (Kourou et al, 2015), fraud detection and automating trading in finance (Arora et al, 2020), as well as the potential of personalized learning experiences in education, tailored to individual students (Luckin et al, 2016). The rapid pace of AI technology demands that we engage these issues more carefully. Here are

just three questions for us to consider. 1. Where does the sense of dread about AI come from? If we look at the history of emerging technologies, should we really be worried? 2. Why has the notion of AI-driven 'post humanity', with our selfhood divided between humans and machines, struck a chord? And why does it resonate with both outrage and yearning? 3. As AI systems mediate more of our lives, society needs to resolve a host of new ethical dilemmas, including algorithmic bias, data privacy and the future of work. How can we, as rational citizens and forward-looking democratic societies, move forward with this transformative technology in the right way?

History of AI

There's a rollercoaster feeling about the story of Artificial Intelligence (AI). Researchers talk about the 'spring', the 'mists of winter', and even the 'winter of doom' in relation to the maturity of certain AI techniques or applications. The key events are well-documented: the term 'Artificial Intelligence' was coined in a conference held at the Dartmouth College in New Hampshire in 1956 by American pioneers such as John McCarthy and Marvin Minsky, who laid the foundations for AI research (Miron & Annoni, 2020). Shortly after, in what authors have described as 'the spring' of AI, researchers developed 'foundational algorithms' and the 'first programs capable of problem-solving and playing games like chess' (ibid, p4). Arthur Samuel's checkers program, which could 'learn from experience', became a media sensation back in 1960. But by the 1970s, optimism had plummeted, and the first AI winter set in: unmet expectations combined with a retraction of funding and interest (prompted in part by the Lighthill report) that lowered research activity to near-zero (Lighthill, 1973). Various early AI applications were similarly criticized, especially those undertaken by the government within the machine translation domain (as detailed in the now-famous ALPAC report, which stood for Automatic Language Processing Advisory Committee (ALPAC, 1966)). The next spring came in the 1980s, when expert systems allowed for a more targeted and reproducible approach to how AI (either new-generation or retrofit) could be 'sold' as AI. The second stage was dominated by machine learning techniques, beginning in the late 1990s and continuing into the early 2000s. The 2010s saw an increased use of deep learning in artificial intelligence, where systems can go through a huge number of examples and learn how to do difficult tasks with remarkable accuracy (LeCun et al, 2015). By example, deep neural networks were used to recognize images and do natural language processing, and other problems. Fast-forward to now, when AI is evolving even more, with growing research into the ethical implications of what may be called Superintelligent AI – that is, systems that might surpass human cognitive capabilities (Bostrom, 2014). The history of AI shapes how we think about its future. For instance, now that machines are able to grasp complex ideas they are also gaining the ability to manipulate, mimic and become glitches themselves, resulting in mischaracterization and misunderstanding. The more we want to understand about AI today, the more important it is to know something about the history.

Types of AI

Artificial intelligence (AI) can be divided into the following distinct subtypes, primarily with respect to the range of tasks they are capable of performing:

- Narrow AI: this is the most common form of AI, also known as Weak AI or Artificial Narrow Intelligence (ANI). Narrow AI can perform specific tasks but only within a restricted domain, such as voice or image recognition, or recommendation systems. For example, a digital assistant like Siri or Alexa is an example of a Narrow AI that easily breaks down commands into machine-processable instructions, but also lacks a true understanding of human conversations (Marr, 2023).
- General AI: in contrast to Narrow AI, General AI (AGI), often referred to as Strong AI, aims to replicate human cognitive abilities across a wide range of tasks. While AGI remains largely theoretical and has not yet been realized, it represents a significant goal for AI research (Bostrom, 2014).
- Superintelligent AI: discussions around Superintelligent AI explore the potential future development of systems that could surpass human intelligence in all domains (Bostrom, 2014).

An increasing number of narrow AI applications are everywhere that work is being processed and – using certain methods of automation and big data analysis – are used to support and simplify decision making. Whether for the delivery of health services, the running of a bank, the selection of TV shows or the

functioning of a computer, narrowly focused AI decreases the time a person must spend deciding on something. For instance, the algorithms used by Netflix and Amazon to steer you towards new content are meant to leverage your eliminated options that favour only the type of content you like because they know your preferences. (DataCamp, 2023). Narrow AI is also used for diagnostics and predictive analytics in healthcare, where efficient processing of large datasets can be crucial for the detection of illnesses (Kourou et al., 2015). Although the epistemic and adaptive limitations of Narrow AI mean there are clear theoretical obstacles to its development as an autonomous, General AI, it continues to drive innovation throughout a wide range of industries.

Key Terminologies

The key terminologies of Artificial Intelligence lay out the basic concepts that any educated person should be familiar with.

- Machine Learning (ML) is a major subfield of AI, defined as learning from experience in a way that will allow it to do something new, without being explicitly programmed – that is, received by learning (Jordan & Mitchell, 2015). Further subdivisions include Supervised Learning, where data with labels on target variables is presented to models, which learn to predict outcomes (Hastie, Tibshirani, Friedman, 2009); and Unsupervised Learning, which deals with unsupervised or unlabelled learning, where shallow transfer learning occurs: 'An entity learns by distilling a pattern from a sea of data' (Hastie, Tibshirani, Friedman, 2009).
- Deep Learning (DL) is a subset of machine learning that involves neural networks which use multiple layers and attempts to learn features from data. Deep learning has had a notable impact on computer tasks such as image and speech recognition, all through the ability to automatically extract features from the form of raw data which is used as input (LeCun, Bengio, & Haffner, 2015).
- Reinforcement Learning (RL) is another key domain which can be thought of as training algorithms to take actions so as to maximize some cumulative reward (eg, money) with the least negative impact in the environment, i.e., RL involves learning how to maximize rewards by trial and error in a dynamic environment (Sutton & Barto, 2018).
- Computer Vision (CV): is a subfield of Artificial Intelligence (AI), enabling machines to interpret visual information from the world, essentially replicating the functionality of the human eye (Szeliski R. 2010). It has wide applications for making decisions in numerous areas, including healthcare (eg, to diagnose through medical images, as described by Vij and Arora (2021), in the automotive industry (eg, so that autonomous vehicles can see and detect objects around them, as detailed by Pendleton et al (2017), and for increasing security through the use of facial recognition technologies (Wang and Deng 2020). As the field advances, its use in an increasing array of applications can be expected.

Understanding these terminologies provides a foundation for exploring the broader implications of AI technologies.

Ethical Considerations in AI

1. Bias in AI Models: Algorithmic bias is a significant ethical concern in AI, leading to discriminatory outcomes based on race, gender, or socioeconomic status. This bias often arises from unrepresentative training data or flawed algorithms that reflect historical inequalities (Miasato & Silva, 2019). For example, résumé-review algorithms used for recruiting earned significantly higher ratings overall for men than women when the training data used was likewise dominated by men (Dastin 2018). Bias can translate to discrimination, and discrimination only exacerbates existing sociodemographic disparities (bias in AI: a brief history of discrimination).
2. Job Displacement: As AI technologies become adopted more broadly in the workplace, some incumbent workers will be displaced from the workforce as machines inevitably perform an ever greater segment of work previously done by humans. In some parts of the economy – such as manufacturing and customer service, for example – hundreds of millions of currently existing jobs could be wiped out by technology (Bessen, 2019) – and reconnecting these displaced workers with

the economy again will require greater investments in reskilling/upskilling (Brynjolfsson McAfee, 2014).

3. AI and privacy: Ironically, the vast volumes of data that AI systems necessitate may be the most detrimental to one's privacy. Everyday surveillance technologies can compromise privacy rights and turn society into a giant brouhaha (Zuboff, 2019). Concerns about the ethics of surveillance lead to data protection measures as a means for protecting individual privacy.
4. Accountability and Transparency: AI decision-making processes are deemed 'black boxes', which means that we do not always know how decisions have been made, and this lack of transparency can complicate liability if algorithms make harmful outcomes. Bridled by clearly defined accountability mechanisms, responsible AI becomes a possibility. Accountability connects to transparency.
5. Autonomous Systems: like military drones and police tools, autonomous systems raise questions about ethics and AI misuse, as well as questions about delegating ethical decisions to machines. These applications would be more likely to be ethical if bound by rules of conduct already.

Future Impact

From diagnosing diseases to driving cars, AI is on course to disrupt industries and economies, enabling as well as causing profound societal shifts. Advanced diagnostic tools will lower rates on missed, delayed, and erroneous diagnoses as well as be matched with tailored treatment plans facilitated by AI (Makridakis, 2017). The advent of fully autonomous vehicles and intelligent traffic-management schemes promise safer and more efficient transportation (Chen et al., 2013), but also threatens mass unemployment (Arntz et al., 2019). As more AI develops, reskilling and retraining will be ever more essential to adjust the workforce to changing demands (Bessen, 2019). Moreover, ethical concerns about the development of AI may soon become central to ensuring that AI technologies that bring us significant benefits are also applied in ways that are fair, address issues related to biases and privacy, and so on (Zuboff, 2019). In conclusion, as AI becomes part of everyday life, we will celebrate advances as we address its challenges – responsibly.

Conclusion

AI technologies bear huge opportunities and challenges we should consider carefully. The AI technologies that are being developed constantly would raise efficiency, help to make better decisions, apply to various industries, and even reach fields like healthcare, transportation, and finance in the future. However, apart from the benefits that AI can bring to different group of population, we should and could not ignore the ethical challenge that AI poses. Players of AI technologies should attach some ethical concerns like bias, data privacy, accountability and the risk of displacing current individuals' jobs etc. When we go into the Internet Era, re-skilling and education of current workforce is always a consideration to overcome the challenges of new technology as a general trend. The same trend will keep happening, especially for the AI technologies. The key would be how we manage to design a system in a way that the AI technologies can really help to raise overall productivity and efficiency and benefit all the groups of people. Policymakers and others should closely work with technologists and the public more frequently to prevent any negative impact to the society and make the potential AI technologies being applied more transparently and beneficially to everyone.

References

Adadi, A. (2021). A survey on data-efficient algorithms in big data era. Journal of Big Data, 8(1). https://doi.org/10.1186/s40537-021-00419-9

ALPAC. (1966). Language and machines: Computers in translation and linguistics. U.S. Government Printing Office.

Arntz, M., Gregory, T., & Zierahn, U. (2019). The Risk of Automation for Jobs in OECD Countries: A Comparative Analysis. OECD Social, Employment and Migration Working Papers, No. 202.

Arora, A., Saini, R., & Gupta, S. (2020). Artificial intelligence in finance: A review of applications and challenges. Journal of Banking and Finance, 113, 105751.

Bessen, J. E. (2019). AI and jobs: The role of demand. NBER Working Paper No. 24235. Retrieved September 17, 2024 from https://www.nber.org/papers/w24235

Bostrom, N. (2014). Superintelligence: Paths, dangers, strategies. Oxford University Press.

Brynjolfsson, E., & McAfee, A. (2014). The second machine age: Work, progress, and prosperity in a time of brilliant technologies. W.W. Norton & Company.

Chen, N., Christensen, L., Gallagher, K., Mate, R., & Rafert, G. (2013). Global Economic Impacts Associated with Artificial Intelligence.

Dastin, J. (2018). Amazon scraps secret AI recruiting tool that showed bias against women. Reuters. Retrieved from https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G

DataCamp. (2023). What is Narrow AI? Retrieved from https://www.datacamp.com/blog/what-is-narrow-ai

Esteva, A., Kuprel, B., Novoa, R. A., et al. (2019). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115-118.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction (2nd ed.). Springer.

Jackson, K. (2021). The impact of algorithmic bias in recruitment. Journal of Business Ethics, 175(2), 345-360.

Jiang, F., Jiang, Y., Zhi, H., et al. (2017). Artificial intelligence in healthcare: Anticipating challenges to ethics, privacy and bias. Health Informatics Journal, 23(4), 305-313.

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349(6245), 255-260.

Jurafsky, D., & Martin, J. H. (2021). Speech and language processing (3rd ed.). Pearson.

Kourou, K., Exarchos, T. P., Emmanouilidis, C., & Koutkias, V. (2015). Machine learning applications in cancer prognosis and prediction. Computational and Structural Biotechnology Journal, 13, 8-17.

LeCun, Y., Bengio, Y., & Haffner, P. (2015). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324.

Lighthill, J. (1973). Artificial intelligence: A general survey. In Artificial Intelligence: A paper on its current status and future prospects (pp. 1-16). British Government.

Lin, P., Bekey, G., & Abney, K. (2012). Autonomous military robotics: Risk, ethics, and design. In Proceedings of the IEEE International Symposium on Ethics in Engineering Science and Technology.

Lipton, Z. C. (2016). The mythos of model interpretability. Communications of the ACM, 59(10), 36-43.

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). Intelligence unleashed: An argument for AI in education. Pearson Education.

Makridakis, P. S. (2017). The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms. ScienceDirect.

Marr, B. (2023). What is Weak (Narrow) AI? Here Are 8 Practical Examples. Retrieved from https://www.linkedin.com/pulse/what-weak-narrow-ai-here-8-practical-examples-bernard-marr

Miasato, M., & Silva, M. (2019). Ethics and discrimination in artificial intelligence-enabled recruitment. Journal of Business Research, 101, 250-258.

Miron, M., & Annoni, A. (2020). Historical evolution of artificial intelligence. AI Watch. Retrieved from https://publications.jrc.ec.europa.eu/repository/bitstream/JRC120469/jrc120469_historical_evolution_of_ai-v1.1.pdf

Mnih, V., Kavukcuoglu, K., Silver, D., et al. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540), 529-533.

Ogbonnia, O. O. (2017). Artificial intelligence as an imperative to the effective delivery of instructional materials' content. International Journal of Advanced Research and Publications, 1(5), 323-327.

Pendleton, S., Andersen, H., Du, X., Shen, X., Meghjani, M., Eng, Y., Rus, D., & Ang, M. (2017). Perception, Planning, Control, and Coordination for Autonomous Vehicles. *Machines*, *5*(1), 6

Samuel, A. L. (1960). Some studies in machine learning using the game of checkers. IBM Journal of Research and Development, 3(3), 210-229.

Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction (2nd ed.). MIT Press.

Szeliski, R. (2010). Computer vision: Algorithms and applications. Springer.

Topol, E. J. (2019). Deep medicine: How artificial intelligence can make healthcare human again. Basic Books.

Vij, R., & Arora, S. (2021). Computer Vision with Deep Learning Techniques for Neurodegenerative Diseases Analysis Using Neuroimaging: A Survey. In *Advances in intelligent systems and computing* (pp. 179–189).

Wang, M., & Deng, W. (2020). Deep face recognition: A survey. *Neurocomputing*, *429*, 215–244.

Zuboff, S. (2019). The age of surveillance capitalism: The fight for a human future at the new frontier of power. PublicAffairs.