International Journal of Electronics and Microcircuits

E-ISSN: 2708-4507
P-ISSN: 2708-4493
IJEM 2022; 2(2): 08-12

www.microcircuitsjournal.com
Received: 04-05-2022
Accepted: 05-06-2022

Samridh Anand Paatni
School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, Tamil Nadu, India

Correspondence
Samridh Anand Paatni
School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, Tamil Nadu, India

AI safety: Public attitudes and necessity for next generation secure communication

Samridh Anand Paatni

Abstract
AI is a promising, upcoming field with applications in every aspect of society. All major industries are looking to use AI in the near future, especially communication. As the fields 5G communication and IoT mature, an increasing attention is placed on AI in managing them. With rapid development and leaps in the field every year, the impact of AI on society has also gained huge attention. This has led to a growing field of research concerned with ensuring that AI technologies are safe and are helpful to society rather than detrimental.
This paper discusses the potential impact of AI on next generation communication, the necessity of the field of AI safety and changing public attitudes toward it. The interaction of human communication and AI systems is also analyzed from a security point of view. Finally, this paper discusses what can be done to change public, academic and commercial attitudes towards AI safety.

Keywords: AI, AI safety, IoT, 5G, next generation communication

1. Introduction
Artificial Intelligence or AI is a relatively new field of computer science. According to John McCarthy, "It is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable" [1]. The ultimate goal of AI research and development is to create machines which can think like humans and to use them to solve human problems.
Communications technologies have played an important part in human development and as they advance, the use of AI systems in communication is gaining more prominence. This paper aims to analyze the interaction between AI and next generation communications technologies.

2. Overview of AI and AI Safety
2.1 Artificial Intelligence
AI is the ability of machines to perform tasks that nor- mally require human intelligence [2]. Humans have natural intelligence and AI researchers aim to create machines with intelligence. Intelligence is the computational part of the ability to achieve goals in the world [1]. Thus, an AI system must be able to perceive its environment, process the data collected and use 'intelligence' to solve a problem.
AI systems learn using 'Machine learning' which is the name for a set of algorithms which mimic human learning by improving over time by collecting and processing data.
Machine learning can be broadly classified into three classes [3]:

1. Supervised Learning: Supervised learning matches in- puts to already-given outputs. The answers are provided for training and the system learns by calculating a 'cost' and optimizing itself. In supervised learning, the goal is often to get the computer to learn a classification system that we have created [3].
2. Unsupervised Learning: Unsupervised learning doesn't have pre-defined outputs. The system either works by a 'reward-function' or essentially, by trial and error [3].
3. Reinforcement Learning: Reinforcement learning works by interactions between an agent and the environment.

AI Safety
AI safety is the field of research concerned with ensuring that the goals of an AI system matches the goals of humanity.

As AI systems get increasingly intelligent with increasing advancement in AI research, concerns about the dangers of AI have received greater attention which has led to increased work in AI safety [4].

The dangers of AI systems can be manifold including danger to individuals, organizations and society [5]. These dangers may stem from any of the steps of development, learning and usage. These risks range from erroneous results to potential risk to humanity. Some specific risks identified by Cheatham, Javanmardian and Samandari [5] are:

1) Risks to Individuals
- Physical safety
- Privacy and reputation
- Digital safety
- Financial health
- Equity and fair treatment

2) Risks to Organizations
- Financial performance
- Non-financial performance
- Legal and compliance
- Reputational integrity

3) Risks to Society
- National security
- Economic stability
- Political stability
- Infrastructure integrity

AI systems learn using what data they were fed thus, their results are only as good as the data they were trained on [6]. Thus, any bias in training will produce a biased result.

A CSET report [6] specifies the following core principles of AI systems for safety:

1) **Robustness:** Making sure that AI systems can work under a wide range of conditions safely and accurately.
2) **Assurance:** Making sure that humans can understand how the system behaves. The developers of an AI system should be able to interpret the results and tell why the they are so, intended or unintended. The current 'black- box' models do not satisfy this need.
3) **Specification:** Making sure that the goals of an AI align with the goals of humans. 'Misspecification' can lead to adverse effects as the results the system produces are not the intended results. Many present-day systems have showed the effects of not meeting this condition, for example the recommendation algorithms of Youtube led people to extremist content in an effort to drive engagement [6].

Another risk is the use of AI systems by bad actors and the weaponisation of AI. Many applications of AI, for example deep-fakes can help in spreading misinformation and propaganda, with AI serving as a tool for authoritarian governments or rogue states.

The most distant but existential risk is of a super-intelligent AI which has surpassed human intelligence and cannot be controlled by humans. This is known as the control problem. This puts a new perspective on AI research, with it being regarded as "bomb that is capable of completely destroying our planet" [7]. These issues illustrate the importance of AI safety as a field.

3. Next Generation Communication

Communications technologies are evolving at a fast pace with many new technologies being on the horizon. Technologies like 5G communication and the Internet of Things (IoT) are two technologies with a huge potential impact on the interconnection of devices. 5G is the newest global wireless standard designed to connect many kinds of devices. It is meant to deliver higher peak data speeds, ultra-low latency, more reliability, massive network capacity, increased availability and a more uniform user experience to more users [8]. According to Oracle, "The Internet of Things (IoT) describes the network of physical objects—"things"—that are embedded with sensors, software, and other technologies for the purpose of connecting and exchanging data with other devices and systems over the internet. These devices range from ordinary household objects to sophisticated industrial tools" [9]. The low latency, high speed networks of 5G and increased interconnectedness of devices creates a lot of new opportunities for technological innovation. However, it also raises concerns of privacy and security. There are many security concerns like privacy protection, authorization and scalability among others [10].

4. AI and Next Generation Communications

Artificial Intelligence is being used in both 5G networks and IoT development.

4.1.1 Overview

AI systems are going to be instrumental in next generation communication, as the complexity of size, volume and speed of data increases, we will need increasingly efficient ways of organizing and making sense of it. The increasing amount of data gives us an opportunity to learn more about the world by using AI and machine learning systems to make sense of otherwise meaningless and unstructured data in large volumes.

4.2 AI for 5G Networks

5G networks allow for data speed and bandwidths which can open a new age of data transfer. According to X. You, C. Zhang, X. Tan *et al.* [11], AI are useful in 5G communication for:

1) Network resource allocation: OFDM resource block allocation in 5G is more complex than before and this can be efficiently done by 'Q-learning', a machine learning technique part of 'Reinforcement learning'.
2) SON: Self-organizing networks are denser than ever before in 5G and AI systems have huge potential in the area by a unique combination of supervised and unsupervised learning [11].
3) Uniform 5G accelerators: CNNs or Convolutional Neural Networks can be used to create better 'belief propogations algorithms'.
4) Optimization of end-to-end physical layer communication: For dense 5G networks and in scenarios that are too complicated to model, autoencoders can be a good solution to 'learn' scenarios and optimize performance.

4.3 AI for IoT

Iot systems consist of a myriad of sensors and devices collecting data. This data can be effectively analyzed using AI. The synergy of a network of devices, connected using a

high-speed data network with the computational power can be revolutionary in many fields. Smart home systems can be improved by using AI to deter-mine how power consumption varies in a house and how can a home be more efficient; smart home security systems can benefit from AI which can detect patterns of unusual behaviour and even contact the authorities if needed; research labs can use a multitude of sensors, all feeding data into a single system which uses AI to parse data in real time. The applications are endless. Some interesting future remarks of AI for IoT have been discussed in [24-40]. We suggest readers/ future researchers to refer these articles [24-40] to know popular issues in current era and suggested countermeasures for the raised issues. Note that such issues may be helpful for their future research to work on further as problem.

4.4 The Need of AI Safety

We have seen that AI systems have applications in many fields and are being developed rapidly, and AI safety should be an integral part of the process. AI is a tool, and a powerful one at that, we must make sure that such a tool is handled responsibly. AI in the hands of tyrannical government nefarious groups or misinformation spreading groups can be a big threat to safety and democracy. The IoT combined with AI systems poses a new threat. As many outdoor IoT devices such as cheap sensors with lesser protections are very vulnerable to attacks [12]. Furthermore, "AI enabled crimes" are a rising threat, a study by Caldwell, Andrews, Tanay, *et al.* [13] in the Crime Science journal classified AI driven crimes based on their harm and how easy they are to defeat. Their findings indicate that audio/video impersonation, phishing attacks, false news stories, using driver-less vehicles as weapons and large-scale blackmail are the biggest threats in AI-driven crimes. These are hard to defeat and can be disastrous in the wrong hands. On the other hand, more minor and easier to deal with crimes like forgery, using AI to trick AI systems, and learning based data attacks are also going to be present. Apart from malicious uses of AI, errors or biases in training data can also have a huge impact on results. As training data is sources from the real world, they can perpetuate real world biases, for example, many AI systems have reproduced the racism in the US police system due to biased data [14]. Thus, AI safety is a necessary field, needed to ensure how and why such powerful tools are built and used.

5. Attitudes towards AI Safety

As we have seen, AI safety is essential in making sure that AI systems are developed securely and their applications are handled responsibly. This section explores how different groups view the field and how their attitudes shape the future of both AI research and AI safety.

5.1 Academic Attitudes

The field of AI safety has changed a lot in recent years. According to a qualitative study [4], AI safety research has boomed with less than a 100 papers published on the topic in 2010 compared to more than 500 papers in 2020. It is clear that academia has understood the need of the field and are doing more and more work on it, however the responsibility of shedding mainstream light on the topic also falls on them. For progress to happen, researchers need to change the views on AI safety of the public, the businesses that are eventually going to use AI and the government, which is a source of money for research and can create laws to enforce standards and regulations.

5.2 Public Attitudes

Current studies point to a generally favourable attitude on AI by the public [15], although with some outliers, particularly in the US, where about 47% of people [16]. As people are getting more aware about AI in general, more outreach about AI safety is necessary. As the public is the one electing officials, it is imperative that AI safety is at the back of their minds as with increasing usage of AI in all aspects of life, they will be the final judges on whether policies on regulation and standards are put into effect. Public attitudes can be a catalyst for broad change, and thus effect businesses and governments. Better media outreach to the next generation by trusted sources can go a long way to improve the knowledge of the public on the subject. Online media can be full of incorrect or malicious information, more authoritative voices on the subject could also increase the awareness of the risks of AI and what should be done to avoid or mitigate them.

5.3 Commercial Attitudes

Private ventures into AI research are going to be critical in the future. As more and more businesses are using AI and machine learning for their products, private companies are often at the forefront of the applications of AI, if not pure research. Surveys show that businesses are the among the most exited groups for AI safety [16]. Thus, it is important that the private sector understand the importance of AI safety as businesses and lobbying groups have a huge impact on government decisions. Some of the biggest proponents of AI safety are business owners, like Elon Musk, who was the founder of Open AI a non-profit AI research company and is a proponent of AI regulation [18]. Although not perfect, such support is vital for the field. Private companies have the capital for big strides in both AI safety, research and applications, making their support necessary. However, there is a need to ensure that a blind profit-motive does not super cede necessary research.

5.4 Political Attitudes

As we have seen, regulations on AI research could be a key part of AI safety. This is an opinion shared by many academic and industry experts [18-19]. Recent developments, like the Request of Information by the US financial regulators of AI use [20] and FTC guidelines on "truth, fairness and equity" in commercial uses of AI [21], have shown that governments are starting to take notice of developments in AI and AI safety. Many world governments have shown interest in the regulation of AI including Canada [22], China, The United Kingdom [23], The European Union and The United States. This has set a precedence that when the time comes, governments won't shy away from stepping in and dictating how the development of AI is handled. For example:

1) In the US, a report by the Defense Innovation Board has made recommendations on the 'ethical use' of AI by the Department of Defense, aiming to ensure the following principles:

a) **Responsible:** Proper judgment by the developers, deplorers and users.
b) **Equitable:** Avoiding unintended biases.
c) **Traceable:** The developers should have an appropriate

understanding of the technology.

d) Reliable: A well-defined domain of use and results.

e) Governable: Avoid unintended harm or disruption.

2) In the European Union, a coordinated plan [24] has identified the following fields to be at a 'high-risk' by AI involvements:

a) Critical Infrastructure (Systems which place human life at risk)
b) Education
c) Safety Components
d) Essential Services
e) Law Enforcement
f) Migration, Asylum and Border Control Manage mint
g) Administration of Justice
h) Democratic Process

The plan also specifies the following obligations for these high-risk systems

a) Adequate risk assessment and mitigation systems
b) High quality datasets
c) Proper logging of activity
d) Detailed documentation
e) Clear information to the user
f) Appropriate human oversight
g) A high level of robustness, security and accuracy

These examples illustrate the kind of views governments nee d to have for proper regulation of AI and responsible AI research. Thus, it is clear that governments, though slow are slowly coming around to view AI as a viable tool and AI safety as an important field.

6. Conclusions

As communications technologies progress, they are using more and more advanced technologies to create a more inter-connected world. AI is one such technology which has a huge potential inside and outside communications. To facilitate the safe and responsible research into AI, AI safety is a field as important as AI development. Communications technologies are advancing at a rapid rate. 5G and IoT are technologies which could possibly change the way the internet, security and local networks work. These, combined with AI can make a huge impact on communications technology, the likes of which were not seen since the invention of the internet. Artificial Intelligence is powerful tool who's responsible and safe usage should be a high priority given the potential scope of its application.

This makes it important that the attitudes towards AI and AI safety of all groups (Public, academic, business and political) are aligned to make a safer, more connected and better world.

7. References

1. McCarthy J. What is artificial intelligence, 2004 Nov.
2. Allen G. Understanding AI technology, Joint Artificial Intelligence Centre Report AD1099286, 2020 Apr.
3. Ayodele TO. Types of Machine Learning Algorithms, New Advances in Machine Learning, Yagang Zhang (Ed.), ISBN: 978-953-307-034-6, InTech,
4. Juric M, Sandic A, Brcic M. AI safety: state of the field through quantitative lens, ar Xiv: 2002.05671, 2020 Apr.
5. Cheatham B, Javanmardian K, Samandari H. Confronting the risks of artificial intelligence, McKinsey Quarterly, 2019 Apr.
6. Rudner TGJ, Toner H. Key concepts in AI safety: an overview, Center for Security and Emerging Technology Issue Brief, 2021 Mar.
7. Ashby M. Ethical Regulators and Super-Ethical Systems, ISSS-2017, 2017 Sep 1.
8. https://www.qualcomm.com/5g/what-is-5g
9. https://www.oracle.com/in/internet-of-things/what-is-iot/
10. Leloglu E. A review of security concerns in internet of things, Journal of Computer and Communications, 5. DOI: 10.4236/jcc.2017.51010
11. You XH, Zhang C, Tan XS, *et al*. AI for 5G: Research Directions and Paradigms Science China Information Sciences, 2018.
12. Xiao L, Wan X, Lu X, Zhang Y, Wu D. IoT Security Techniques Based on Machine Learning: How Do IoT Devices Use AI to Enhance Security?, IEEE Signal Processing Magazine. 2018 Sept;35(5):41-49. DOI: 10.1109/MSP.2018.2825478.
13. Caldwell M, Andrews JTA, Tanay T, *et al*. AI-enabled future crime Crime Science, 2020, 9(14). https://doi.org/10.1186/s40163-020-00123- 8
14. O'Donnel RM. Challenging Racist Predictive Policing Algorithms under the Equal Protection Clause, NYU. Law Review, 2019.
15. Vasiljeva T, Kreituss I, Lulle I. Artificial Intelligence: The Attitude of the Public and Representatives of Various Industries, Journal of Risk and Financial Management, 2021. https:://doi.org/10.3390/jrfm14080339
16. Neudert LM, Knuutila A, Howard PN. Global Attitudes Towards AI, Machine Learning & Automated Decision Making Implications for Involving Artificial Intelligence in Public Service and Good Gover- nance, Oxford Commissions on AI & Good Governance, 2020.
17. Markoff J. Artificial-Intelligence Research Center is Founded by Sili-con Valley Investors, The New York Times, 2020 Sept 30th.
18. Gibbs S. Elon Musk: artificial intelligence is our biggest existential threat, The Guardian, 2015 Oct 30th.
19. Buiten MC. Towards Intelligent Regulation of Artificial Intelligence, European Journal of Risk Regulation. 2019;10(1):41-59.
20. https://www.govinfo.gov/content/pkg/FR-2021-03-31/pdf/2021-06607. pdf
21. https://www.ftc.gov/news-events/blogs/business-blog/2021/04/ aiming-truth-fairness-equity-your-companys-use-ai
22. UNESCO Science Report: The Race Against Time for Smarter Development UNESCO ISBN 978-92-3-100450-6, 2021
23. Leslie D. Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector, 2019. DOI:10.5281/zenodo.3240529.
24. Europe fit for the Digital Age: Commission proposes new rules and actions for excellence and trust in Artificial Intelligence, European Commission, 2021.
25. AI Principles: Recommendations on the Ethical Use of Artificial Intelligence by the Department of Defense, Defense Innovation Board, 2020.
26. Nair MM, Tyagi AK, Sreenath N. The Future with

Industry 4.0 at the Core of Society 5.0: Open Issues, Future Opportunities and Challenges, 2021 International Conference on Computer Communication and Informatics (ICCCI), 2021, 1-7, DOI: 10.1109/ICCCI50826.2021.9402498.
27. Tyagi AK, Fernandez TF, Mishra S, Kumari S. Intelligent Automation Systems at the Core of Industry 4.0. In: Abraham A, Piuri V, Gandhi N, Siarry P, Kaklauskas A, Madureira A. (eds) Intelligent Systems Design and Applications. ISDA 2020. Advances in Intelligent Systems and Computing. Springer, Cham, 2021, 1351. https://doi.org/10.1007/978-3-030-71187-0_1
28. Goyal Deepti, Tyagi Amit. A Look at Top 35 Problems in the Computer Science Field for the Next Decade, 2020. 10.1201/9781003052098-40.
29. Amit Kumar Tyagi, Dr. Meenu Gupta, Aswathy SU, Chetanya Ved. Healthcare Solutions for Smart Era: An Useful Explanation from User's Perspective, in the Book Recent Trends in Blockchain for Information Systems Security and Privacy, CRC Press, 2021.
30. Varsha R, Nair SM, Tyagi AK, Aswathy SU, RadhaKrishnan R. The Future with Advanced Analytics: A Sequential Analysis of the Disruptive Technology's Scope. In: Abraham A., Hanne T., Castillo O., Gandhi N., Nogueira Rios T., Hong TP. (eds) Hybrid Intelligent Systems. HIS 2020. Advances in Intelligent Systems and Computing. Springer, Cham 2021, 1375. https://doi.org/10.1007/978-3-030-73050-5_56
31. Tyagi Amit Kumar, Nair Meghna Manoj, Niladhuri Sreenath, Abraham Ajith. Security, Privacy Research issues in Various Computing Platforms: A Survey and the Road Ahead, Journal of Information Assurance & Security. 2020;15(1):1-16.
32. Madhav AVS, Tyagi AK. The World with Future Technologies (Post-COVID-19): Open Issues, Challenges, and the Road Ahead. In: Tyagi AK, Abraham A, Kaklauskas A. (eds) Intelligent Interactive Multimedia Systems for e-Healthcare Applications. Springer, Singapore, 2022. https://doi.org/10.1007/978-981-16-6542-4_22
33. Mishra S, Tyagi AK. The Role of Machine Learning Techniques in Internet of Things-Based Cloud Applications. In: Pal S., De D., Buyya R. (eds) Artificial Intelligence-based Internet of Things Systems. Internet of Things (Technology, Communications and Computing). Springer, Cham, 2022. https://doi.org/10.1007/978-3-030-87059-1_4
34. Akshara Pramod, Harsh Sankar Naicker, Amit Kumar Tyagi. Machine Learning and Deep Learning: Open Issues and Future Research Directions for Next Ten Years, Book: Computational Analysis and Understanding of Deep Learning for Medical Care: Principles, Methods, and Applications, 2020, Wiley Scrivener, 2020.
35. Shabnam Kumari, Amit Kumar Tyagi, Aswathy SU. The Future of Edge Computing with Blockchain Technology: Possibility of Threats, Opportunities and Challenges, in the Book "Recent Trends in Blockchain for Information Systems Security and Privacy, CRC Press, 2021.
36. Tyagi AK. (Ed.). Data Science and Data Analytics: Opportunities and Challenges (1st ed.). Chapman and Hall/CRC, 2021. https://doi.org/10.1201/9781003111290
37. Tyagi AK, Abraham A. (Eds.). Recent Trends in Blockchain for Information Systems Security and Privacy (1st ed.). CRC Press, 2021. https://doi.org/10.1201/9781003139737
38. Tyagi AK, Rekha G, Sreenath N. (Eds.). Opportunities and Challenges for Blockchain Technology in Autonomous Vehicles. IGI Global, 2021. http://doi:10.4018/978-1-7998-3295-9
39. Tyagi AK. (Ed.). Multimedia and Sensory Input for Augmented, Mixed, and Virtual Reality. IGI Global, 2021. http://doi:10.4018/978-1-7998-4703-8
40. Malik S, Bansal R, Tyagi AK. (Eds.). Impact and Role of Digital Technologies in Adolescent Lives. IGI Global, 2022. http://doi:10.4018/978-1-7998-8318-0