“… what’s wrong is that ingrained biases in society have led to unequal outcomes in the workplace, and that isn’t something you can fix with an algorithm.”

Dr. Rumman Chowdhury, Accenture

**Responsible Artificial Intelligence**

**by**

Abimbola Ayeni

Trainees, Tech Talent Academy

Under the Supervision of

Georgina Stanley

(Technical Trainer)

1st February 2022

**Responsible Artificial Intelligence (RAI)**

Artificial Intelligence (AI) is a branch of computer science implemented in machines in form of algorithms driven by data to perform tasks that can simulate human intelligence in terms of reasoning, learning, problem-solving and quick decision making (Datatilsynet, 2018 & Tathagata, 2021). AI systems have impressive predictive prowess in everything. But like any new technology, it has flaws, vulnerabilities and unintended effects (Quest, 2022). As the contribution of AI to humanity has been monumental, its failures have also been equally hilarious (Harshajit, 2019).

***If AI systems are not regulated,*** it can lead to the demise of humanity (Jowallah, 2021). For example, in 2016, bot tricked thousands of Facebook users to unknowingly install malware that allowed remote access to their Facebook accounts (Harold, 2019). In 2017, when Oliver Haberstroh was away from home, Alexa started playing music on its own at 1:50am and turned the volume so up (Harshajit, 2019). The CEO of a UK-based energy firm received a call from his German boss instructing him to transfer €220,000 to a Hungarian supplier. Unfortunately, the boss was a ‘deep fake’ voice generation software (ImmuneWeb, 2022). An Uber self-driving SUV killed a female pedestrian in Tempe, Arizona. Uber discovered that its software decided not to act after the car’s sensors detected her (Medium, 2018). In line with this, a ***Responsible AI framework***, based on the principle to create algorithms that foster good ethics and morals when deployed is needed. The deployed algorithms will be secure, transparent, fair, inclusive and would limit levels of discrimination in our society (Jowallah, 2021).

***When machine learning goes wrong* and result in a fatal error,** the responsibilities lies on the developer, the trainer and the operator (Robbrecht, 2017). The European Union General Data Protection Regulation (GDPR) aimed at providing effective legal safeguards for data-subjects, including rights of transparency, explanation and contestation, for fully automated decisions (Quest, 2022). **Article 22** of the GDPR states that data subjects have the right not to be subject to a decision based solely on automated processing. Therefore, individuals may object to any processing of their data which is carried out without any human involvement. Also, **Articles 13-15** of the GDPR requires organisation planning to feed data into an algorithm, that will then decide, to inform individual that processing of this nature will take place (PWC, 2019 & Mark et. al, 2020). Under the GDPR, the EU’s data protection authorities can **impose fines** of up to up to €20 million (Tessian, 2020).

**Therefore, before companies can develop algorithms**, they must have clearly defined policies and guidelines of what user security really means. They must understand and practice fairness and inclusion and have a policy that addresses accountability when violation occur. The system should be designed in a collaborative team approach which should include software development, individuals, managers, industrial leaders and policy makers (Jowallah, 2021). The inbuilt integrity will ensure that segments of our population are not denied access to services based on their ethnicity. A RAI system would not use student script to project possible crimination in later years or use data to profile specific people based on race, religion, skin colour and immigration status (Jowallah, 2021).

**References**

[1] Alex Salwalker & Vivek Wadhwa (2019): commentary on “A.I. Bias Isn’t the Problem. Our Society Is”, <https://fortune.com/2019/04/14/ai-artificial-intelligence-bias/>

[2] Datatilsynet (2018): “Artificial intelligence and privacy Report”, The Norwegian Data Protection Authority January 2018

<https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf>

[3] Glad You Asked (2021): “ Are we Automating Racism?”, [Glad You Asked  S2 • E2](https://www.youtube.com/playlist?list=PLJ8cMiYb3G5cOFj1VQf8ykNOI0ptuHybc)

<https://www.youtube.com/watch?app=desktop&v=Ok5sKLXqynQ>

[4] Harold Kilpatrick (2019): “How Artificial Intelligence Can Be Used Maliciously in Cybersecurity”, By Harold Kilpatrick in Guest Articles, Jan 30th 2019.

<https://www.datapine.com/blog/ai-in-cybersecurity/>

[5] Harshajit Sarmah (2019): “Top 5 Epic Artificial Intelligence Fails”, Analytic India Magazine Pvt Ltd, <https://analyticsindiamag.com/top-5-epic-artificial-intelligence-fails>

[6] ImmuniWeb (2022): ImmuniWebSecurity Blog,

<https://www.immuniweb.com/blog/top-10-failures-of-ai.html>

[7] Jowallah (2021): “[Building Our World With Responsible Artificial Intelligence (AI)](https://www.youtube.com/watch?v=mBlI3TWgbjM)”,

[The Learning Journey with Dr.](https://www.youtube.com/channel/UCV2pwUTa53PLPMzIepBURsg" \o "See channel The Learning Journey with Dr. Jowallah on YouTube" \t "_blank) [Jowallah](https://www.youtube.com/channel/UCV2pwUTa53PLPMzIepBURsg" \o "See channel The Learning Journey with Dr. Jowallah on YouTube" \t "_blank)

<https://www.bing.com/videos/search?q=responsible+artificial+intelligence&docid=607995184087116535&mid=B2E8B4056FC576920FB7B2E8B4056FC576920FB7&view=detail&FORM=VIRE>

## [8] Mark Adair, Mason Hayes & Curran (2020): “Ireland: AI and Automated Decisions: GDPR and the Evolving Decision-Making Process”, 23 January 2020

<https://www.mondaq.com/ireland/Privacy/862034/AI-And-Automated-Decisions-GDPR-And-The-Evolving-DecisionMaking-Process>

# [9] Medium (2018): “2018 in Review: 10 AI Failures”,

<https://medium.com/syncedreview/2018-in-review-10-ai-failures-c18faadf5983>

[10] PWC (2019): “[Data Protection](https://pwc.blogs.com/data_protection/): Artificial Intelligence (AI) and the GDPR - Part one

<https://pwc.blogs.com/data_protection/2019/01/artificial-intelligence-ai-and-the-gdpr-part-one.html>

[11] QCS (2020): “[Six Principles of GDPR - GDPR | QCS Blog](https://www.qcs.co.uk/six-principles-of-gdpr/?utm_source=bing&utm_medium=cpc&utm_adgroup=All%20web%20pages%20-%20Google%20Index&utm_campaign=01_England_DSA&matchtype=b&keyword=qcs%20co&network=o&device=c&msclkid=1806bab5299e13c290dadb40ff2c5874)”,

<https://www.qcs.co.uk/six-principles-of-gdpr/?utm_source=bing&utm_medium=cpc&utm_adgroup=All%20web%20pages%20-%20Google%20Index&utm_campaign=01_England_DSA&matchtype=b&keyword=qcs%20co&network=o&device=c&msclkid=1806bab5299e13c290dadb40ff2c5874>

[12] Quest (2022): “How has the law been pushed aside in the age of AI?”, Quest Contributor, University of Birmingham

<https://www.birmingham.ac.uk/research/quest/emerging-frontiers/AI-and-the-law.aspx>

## [13] Robbrecht van Amerongen (2017): “Who is responsible when Artificial Intelligence fails?”

<https://www.linkedin.com/pulse/who-responsibility-when-artificial-intelligence-amerongen/> Published on January 30, 2017

[14] Tathagata Das (2021): “What is Artificial Intelligence (AI)?”

https://tech4fresher.com/what-is-artificial-intelligence-ai/

[15] Tessian (2020): “25 Biggest GDPR Fines So Far (2019, 2020, 2021, 2022)”,

<https://www.tessian.com/blog/biggest-gdpr-fines-2020>

Today, the tech giants deploy A.I. to influence what we see, hear, buy—and even feel. The shift of the burden of discernment from people to machines is rapidly spreading into many other nooks—but under the guise of “optimization.”

It’s time to press pause. Perhaps in the future we can create systems that do an excellent job of utilizing data about our lives while excising bias. Until then, A.I. left unchecked is more of a risk than a benefit to society.

Alex & Vivek (2019)