

## Artificial Intelligence: Responsible AI

Responsible AI is a governance framework that can include details on what data can be collected and used, how models should be evaluated, and how to best deploy and monitor models. The framework can also define who is accountable for any negative outcomes of AI. The key principles of Responsible AI underscore fairness, transparency and explainability, human-centeredness, and privacy and security.

I have read about Amazon recruiting tool that was filtering out candidates unfairly. Amazon had been building software that would automate the process of reviewing job applicants' resumes with the intention of finding the top 5 talents since 2014. It was not until 2015, Amazon's machine learning specialists discovered that their AI-powered recruiting tool was hiring for technical roles (i.e. software developer) in a way that was not gender-neutral.

Amazon had trained their machine learning algorithms on resumes that had been submitted to the company over a 10-year period. The majority of resumes came from men, since this is what was most common in technical roles, and the algorithms learned this pattern and determined women are not good suitors for technical roles.

According to Kalliopi Spyridaki, Chief Privacy Strategist at SAS Europe, the GDPR only covers several key areas specifically pertaining to AI. When AI processes personal data, performs profiling, or makes automated decisions based on personal data and/or that affect the data subject, the GDPR applies. This includes the right to object to personal data processing, the right of access, the right to be forgotten, etc. The GDPR also gives individuals the right to not be subjected to solely automated decision-making **except in certain instances**. Article 22 is the only provision within the GDPR that specifically applies to automated decision-making. However, it only applies in a limited set of circumstances and cannot be considered in isolation. There are forms of processing with legal or legally-significant effects that include medical decisions (e.g., those affecting treatment), algorithmic filtering when hiring, access to education, and access to credit. Therefore, in cases when the structure of the algorithm or the data being processed could lead to life-changing errors, privacy breaches, or instances of discrimination, Article 22 must apply. However, Article 22 does not apply when explicit consent to processing is given, when the processing is necessary to the performance of a contract, or when it is authorized by another Union or member state law. The problem is that existing AI system logic takes automated decisions without user consent. I would suggest that the big companies should make a common effort to overcome these difficulties by designing a corporate framework which sets limitations on AI involvement in critical processes like hiring, employee profiling and business data handling.