In Data Science we process a lot data through AI. With the

GDPR, it is becoming increasingly important to understand

the ethics behind the data that is collected, stored,

processed and evaluated.

Your task is to:

• Find out what Responsible AI is?

Responsible AI is a governance framework that documents how a specific organization is addressing the challenges around artificial intelligence (AI) from both an ethical and legal point of view. Resolving ambiguity for where responsibility lies if something goes wrong is an important driver for responsible AI initiatives.

• Find instances where AI has failed? Or been used

maliciously or incorrectly.

Data analysis and machine learning methods can be used for both benign and malicious purposes. Analytics techniques used to plan marketing campaigns can be used to plan and implement effective regional or targeted spam campaigns. Data freely available on social media platforms can be used to target users or groups with scams, phishing, or disinformation. Data analysis techniques can also be used to perform efficient reconnaissance and develop social engineering strategies against organizations and individuals in order to plan a targeted attack.

The potential impact of combining powerful data analysis techniques with carefully crafted disinformation is huge. Disinformation now exists everywhere on the Internet and remains largely unchecked. The processes required to understand the mechanisms used in organized disinformation campaigns are, in many cases, extremely complex. After news of potential [social media manipulation of opinions during the 2016 US elections](https://www.politico.eu/article/cambridge-analytica-chris-wylie-brexit-trump-britain-data-protection-privacy-facebook/), the [2016 UK referendum on Brexit](https://www.newyorker.com/news/news-desk/new-evidence-emerges-of-steve-bannon-and-cambridge-analyticas-role-in-brexit), and [elections](https://www.voanews.com/africa/cambridge-analytica-played-roles-multiple-african-elections) [across](https://www.newstatesman.com/world/2018/03/cambridge-analytica-facebook-elections-africa-kenya) [Africa](http://privacyinternational.org/feature/954/texas-media-company-hired-trump-created-kenyan-presidents-viral-anonymous-attack), and [Germany](https://www.texasmonthly.com/the-daily-post/how-an-austin-ad-agency-helped-the-alt-right-rise-again-in-germany) many governments are now worried that well-organized disinformation campaigns may target their voters during an upcoming election. Election meddling via social media disinformation is [common](https://www.buzzfeednews.com/article/ryanhatesthis/meet-the-29-year-old-trying-to-become-the-king-of-mexican) in [Latin](https://blog.usejournal.com/social-media-automation-information-warfare-by-the-venezuelan-opposition-9cdb407492f8) [American](https://medium.com/@erin_gallagher/mexico-coordinated-inauthentic-behavior-on-facebook-twitter-a670280d02fc) [countries](https://medium.com/@erin_gallagher/fake-honduran-twitter-the-digital-campaign-against-berta-c%C3%A1ceres-and-copinh-3d1ea62e61ab). However, in the west, disinformation on social media and the Internet is no longer solely focused on altering the course of elections – it is about [creating social divides](https://www.wired.com/story/russia-ira-propaganda-senate-report/), causing confusion, manipulating people into having more extreme views and opinions, and misrepresenting facts and the perceived support that a particular opinion has.

• Implications of when AI fails. There is a specific article in

the GDPR Law that covers this, especially with automated

decision making. (opt in and out options).

The data subject should have the right not to be subject to a decision, which may include a measure, evaluating personal aspects relating to him or her which is based solely on automated processing and which produces legal effects concerning him or her or similarly significantly affects him or her, such as automatic refusal of an online credit application or e-recruiting practices without any human intervention. 2Such processing includes ‘profiling’ that consists of any form of automated processing of personal data evaluating the personal aspects relating to a natural person, in particular to analyse or predict aspects concerning the data subject’s performance at work, economic situation, health, personal preferences or interests, reliability or behaviour, location or movements, where it produces legal effects concerning him or her or similarly significantly affects him or her. 3However, decision-making based on such processing, including profiling, should be allowed where expressly authorised by Union or Member State law to which the controller is subject, including for fraud and tax-evasion monitoring and prevention purposes conducted in accordance with the regulations, standards and recommendations of Union institutions or national oversight bodies and to ensure the security and reliability of a service provided by the controller, or necessary for the entering or performance of a contract between the data subject and a controller, or when the data subject has given his or her explicit consent. 4In any case, such processing should be subject to suitable safeguards, which should include specific information to the data subject and the right to obtain human intervention, to express his or her point of view, to obtain an explanation of the decision reached after such assessment and to challenge the decision. 5Such measure should not concern a child.