Bot Detection on X

Thesis Proposal

giovanni.stilo@univaq.it, andrea.dangelo6@graduate.univaq.it

The training of Generative AI

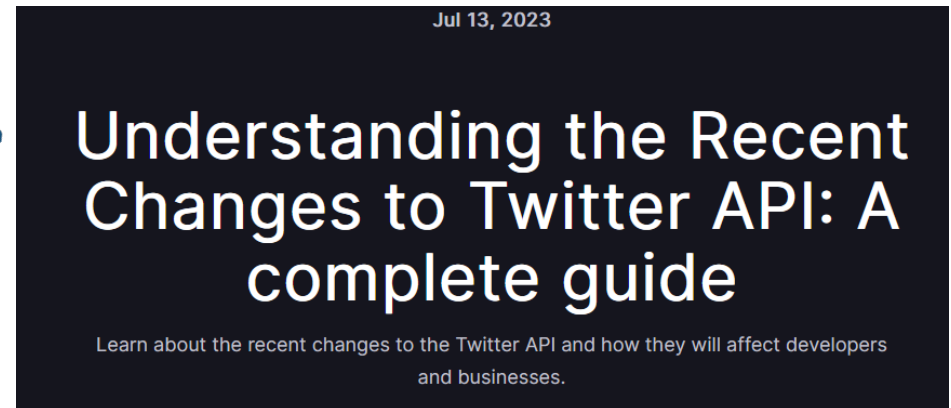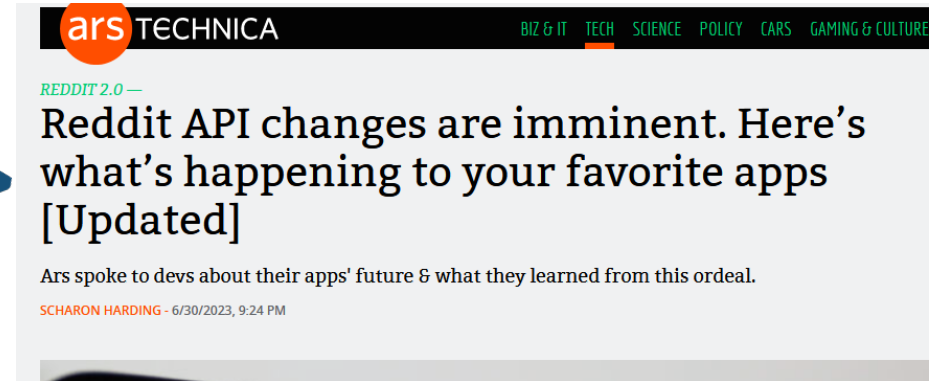**Gen AI Models** are trained with a large amount of unlabeled text

Wikipedia, **Reddit** and **Twitter** are just examples of text sources.

The model will learn from the text, possibly memorizing information.

What are the consequences?

Reddit and Twitter changed their API policies to profit from mass scale access.

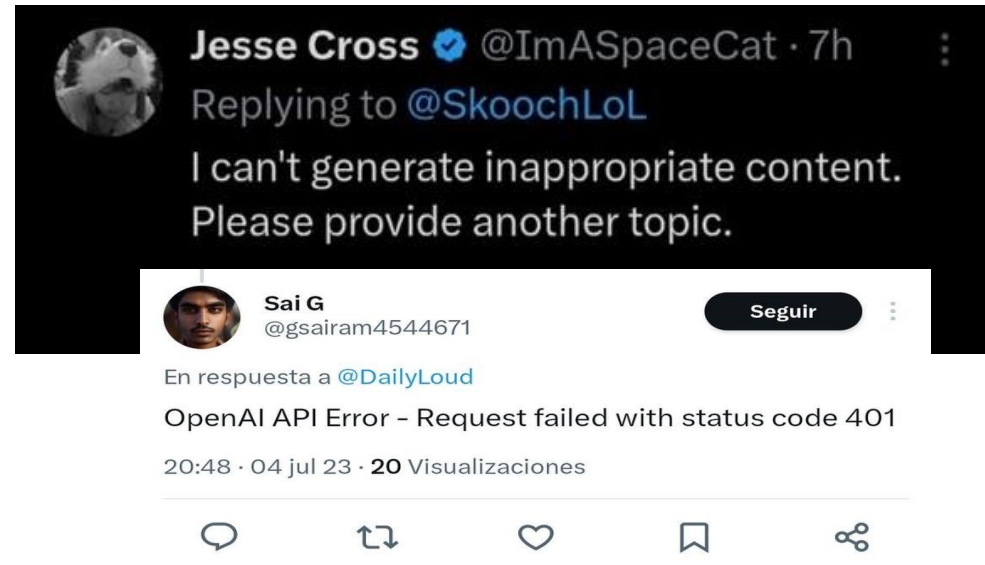AI Bots are taking over the entire Internet.

Theory of Dead Internet

The Theory of Dead Internet states that the Internet is now almost entirely dominated by AI Bots.

The post

The comments

Thesis
Main points

- Study the literature on LLM bots detection

- Apply methods on X (formerly twitter)

- Clustering, Anomaly Detection, Community Detection, Graph Analysis on the results