

BTO - Meeting between Roberto Viola and Open AI, 31 March 2023.

Personal data

Participants:

OpenAI: [REDACTED]

DG CNECT: Roberto Viola, Lucilla Sioli, [REDACTED]

Commercial interests

OpenAI started the meeting explaining GPT4 and showing some demos, for example Sartre talking about fundamental rights and advise for lawyers on opening a chocolate business. The system is now multi-lingual. [REDACTED]

[REDACTED]. Open AI explained how the developer can introduce prompts and make sure that the system does not do harm. (For example the system prompt could be to give no advice that leads to physical harm.) It is the developer's choice what he uses the application for and what classifiers are being put.

[REDACTED]

OpenAI acknowledges that a clear regulatory framework could make sense. However, the possibility of continuous learning was found very important. [REDACTED]

OpenAI supports transparency. However, transparency always needs to be balanced against safety (in particular regarding the training data and model, this would open the door for the creation of malicious systems and the training data is not found critical to understand the capabilities of the system) and trade secrets need to be respected. Sharing the capabilities, limitations and testing results on the other hand is useful.

[REDACTED]

OpenAI emphasised that it is building inherently safe foundational models that are subject to a number of safeguards and moderation.

Roberto Viola thanked OpenAI for the comprehensive information.

Disclaimer: This document was drawn up for internal use under the responsibility of the relevant staff member of DG CONNECT. It solely reflects the author's interpretation of the interventions made. It does not reflect the position of nor does it commit the Commission, DG CONNECT and/or any third party referred to and cannot be quoted as such.