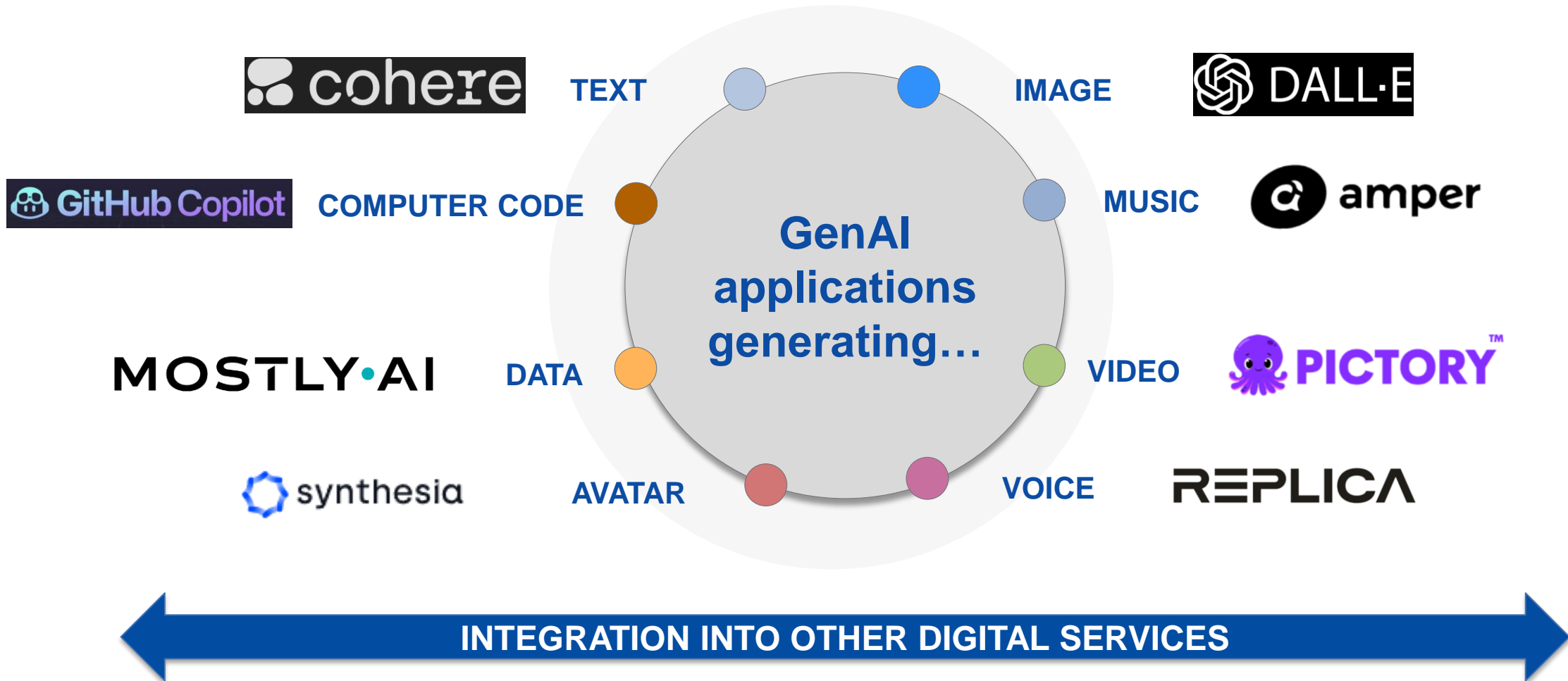


Developing practices to opt-out or license content for GenAI training

European Observatory on Infringements of Intellectual Property Rights
Antoine Aubert
23 April 2024

The GenAI Application landscape



Training of GenAI models

AI model trained by analysing large amount of data

To draw insights from the form of text, images or other media (input) to learn themes, techniques and concept and generate new content with similar characteristics (output).

Source of the training data?

From curated, labelled and sometime opened or public datasets to content available online.

AI Act – ‘Copyright relevant’ provisions

Obligations applying to all General Purpose AI models placed on the EU market:

- Put in place a **policy to respect Union copyright law**, in particular to **identify and respect the opt-out from the ‘text and data mining’ exception**
- Make **publicly available a sufficiently detailed summary** of the data used for the training
 - Comprehensive in its scope in a narrative form, not technically detailed
 - Aim to facilitate parties with legitimate interests, including rights holders, to exercise and enforce their rights under Union law, while balancing interests and trade secrets of providers
 - The AI Office to provide a template for the summary

Developing practices to opt-out or license content for training purposes

robots.txt

TDM Reservation Protocol (TDMRep)
Final Community Group Report 02 February 2024



TDM·AI

**TECHNICAL
STANDARDS**

**REPOSITORY OF
OPTED OUT
REPERTOIRES**

RIGHTSANDAI.COM

**Developing
Opt-out
practices**

**OPT OUT OF
ENTIRE
REPertoire**

**STANDARD
CLAUSES**



Next Step - Forthcoming EUIPO study on GenAI and copyright

≡ KEY ELEMENTS OF THE ANALYSIS ON GenAI INPUT

USE OF CONTENT

Through different methods & at the different phases of GenAI system training including unlearning

SOLUTIONS & PRACTICES

To limit, reserve, support access or license the use of copyright protected works as training materials.

Next Step - Forthcoming EUIPO study on GenAI and copyright

≡ KEY ELEMENTS OF THE ANALYSIS ON OUTPUT

GENERATION OF CONTENT

Through different methods & phases including defining and analysing prompts, refining output, content moderation and validation

SOLUTIONS & PRACTICES

To identify and/or mark AI generated works & prevent regurgitation



www.euipo.europa.eu

 [@EU_IPO](https://twitter.com/EU_IPO)

 [EUIPO](https://www.linkedin.com/company/euipo)

 [EUIPO.EU](https://www.facebook.com/EUIPO.EU)

 [@EUIPO](https://www.instagram.com/EUIPO)

 [EUIPO](https://www.youtube.com/EUIPO)

THANK YOU!