General AI Models in Service of Open Science –
What are the possibilities and limitations of the current EU copyright and data legislation to enable large-scale AI and data science research?

Teresa Nobre, Legal Director
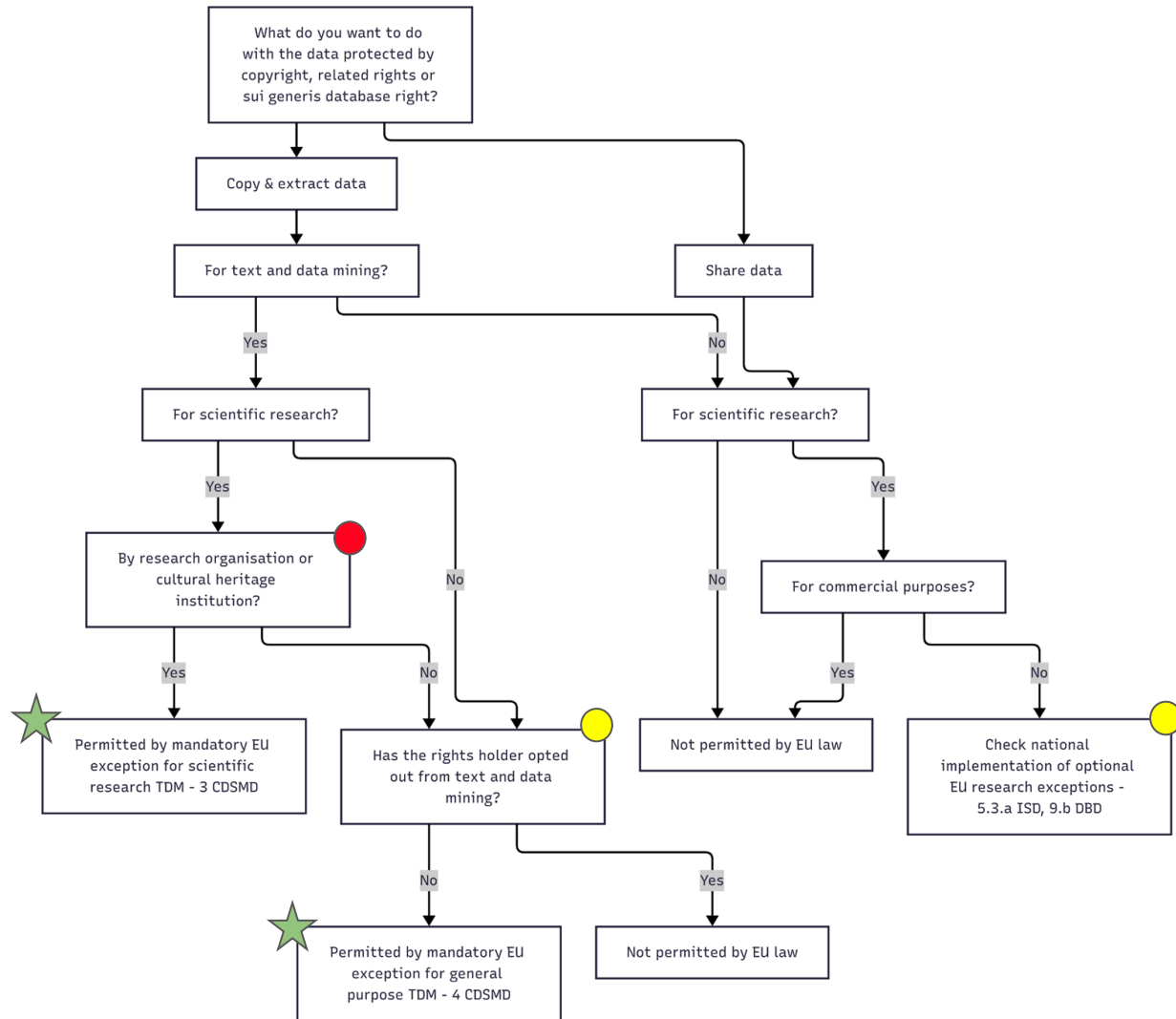
COMMUNIA

AI training data: legal frameworks

- Lawful access
  - Publicly and freely available
  - Open/commercial licenses
  - Data access rights
- Lawful use
  - Public Domain
  - Copyright exceptions (CDSMD, ISD, DBD)
  - Open/commercial licenses

Copyright framework: opportunities

- AI Act confirms that **TDM exceptions cover AI training**
- CDSMD harmonized rules for **copying & extract data for TDM**
  - By research orgs & CHI for scientific research (protected against contracts)
  - By anyone for any purposes (subject to opt-outs)
- Some MS allow **copying & sharing data** for **NC scientific research** (ISD, DBD)
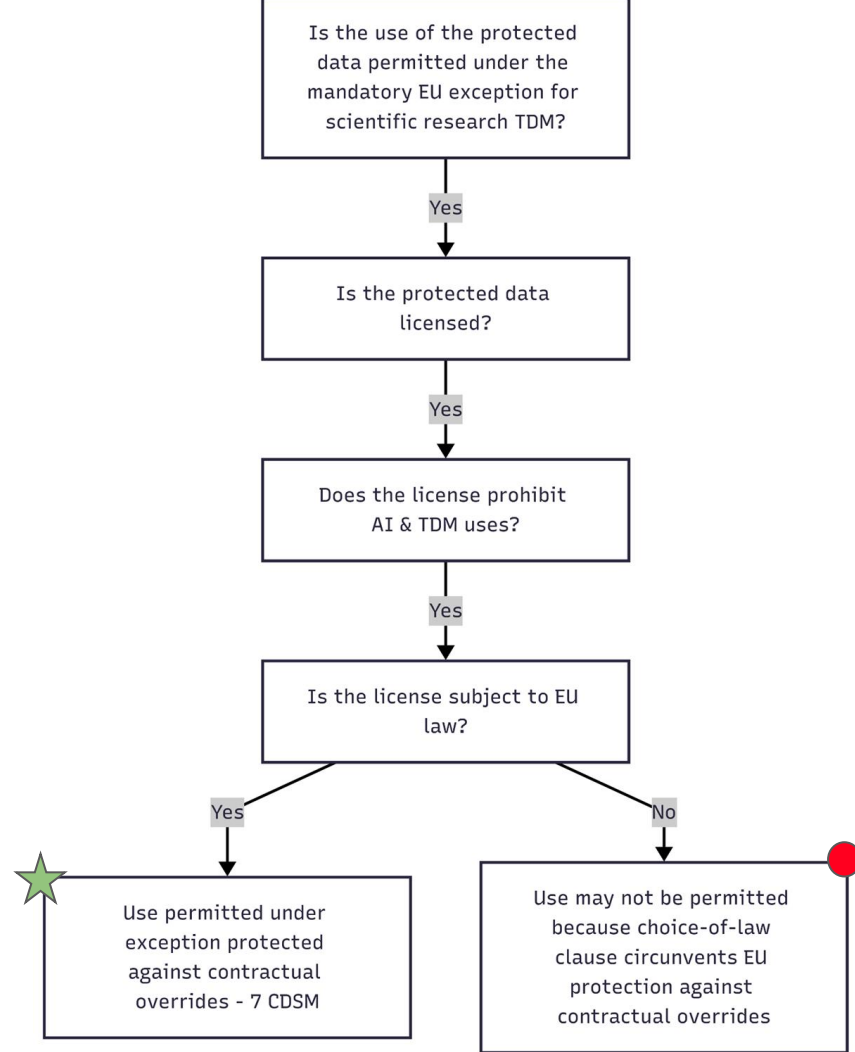
Copyright framework

What do you want to do with the data protected by copyright, related rights or sui generis database right?

Copy & extract data

For text and data mining?
- Yes → For scientific research?
- No → Has the rights holder opted out from text and data mining?

For scientific research? (from Copy & extract data left branch)
- Yes → By research organisation or cultural heritage institution? ●(red)
  - Yes → ★ Permitted by mandatory EU exception for scientific research TDM - 3 CDSMD
  - No → Has the rights holder opted out from text and data mining? ●(yellow)
- No → Has the rights holder opted out from text and data mining?

Has the rights holder opted out from text and data mining?
- No → ★ Permitted by mandatory EU exception for general purpose TDM - 4 CDSMD
- Yes → Not permitted by EU law

Share data

For scientific research?
- No → Not permitted by EU law
- Yes → For commercial purposes?
  - Yes → Not permitted by EU law
  - No → Check national implementation of optional EU research exceptions - 5.3.a ISD, 9.b DBD ●(yellow)

Copyright framework: challenges

- **Copying and sharing data for NC scientific research** not protected across the EU

- **Public-private partnerships** not covered by harmonized exception for scientific research TDM

- **Contractual prohibitions and TPM restrictions** to TDM and AI

- **Making opt-outs effective:** standards; register; transparency; fair remuneration

- **Infringing AI outputs**

Contractual prohibitions to TDM and AI

Thank you!