Text and Data Mining
Do we need to distinguish between commercial and non-commercial use?
Presentation outline

● What is text and data mining?

● Why does it matter to distinguish?

● How to distinguish commercial from non-commercial?

● What is the risk of such distinction?
What is text and data mining?

Data mining = an analytical process that looks for trends and patterns in data sets that reveal new insights.

The outcome of TDM activities results in pieces of information previously unknown and potentially useful.
What is text and data mining?

Examples:

- [www.contentmine.org](http://www.contentmine.org) - Develop software for mining, to liberate 100,000,000 facts from scientific literature

- [www.sparrho.com](http://www.sparrho.com) - Search engine to recommend scientific information

Why is it important now?

The distinction between commercial and non-commercial is being pushed in future EU legislative proposals.
Why does it matter how it is qualified?
What is commercial? What is non-commercial?

No agreed definition. Some hints:

- Who's doing the TDM activities?
- What's the motivation/aim of the TDM activities?
- What's the source of the input material for TDM?
- How is TDM output used / where it is published?
- Who owns the derivative data?
What is commercial?
What is non-commercial?

Nota bene!
These are strictly legal concepts.

While conducting their work, researchers would not typically ask themselves if what they do is commercial or non-commercial.

=> it constitutes an additional constraint
Practical examples

• Is teaching a commercial use? (Court case in Germany)

• Is a research project that's funded through public-private partnerships and the outcome of which is published in for-profit journals commercial or non-commercial?

...and ... who gets to decide?
The risk

A different legislative regime for commercial vs. non-commercial uses respectively

... when in practice there are almost non-existent 100% non-commercial uses

**Result**: risk of missing opportunities for innovation
Thank you!

Diana Cocoru
Contact: diana@openforumeurope.org
www.openforumeurope.org