The Challenges of Compensating Owners of Copyright to Works in Artificial Intelligence's ("AI") Training Data
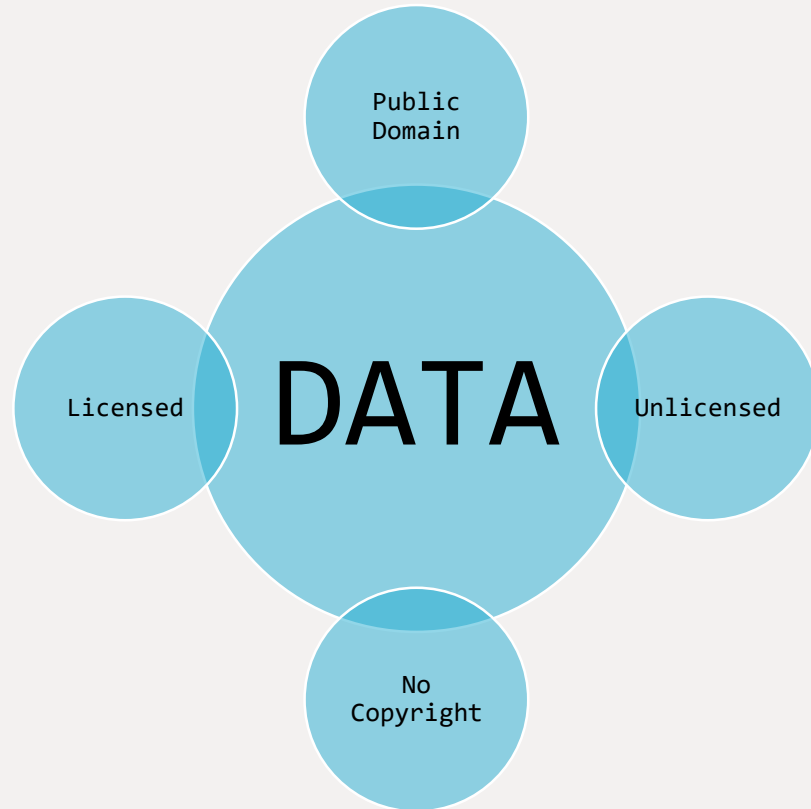
Zeynep Ülkü Kahveci
LL.M.

Training Data

Q:

How do we classify training data of an artificial intelligence system during the lifecycle thereof?

Exceptions to the Copyright Regime

The US

❑ Copyright Act

❑ Section 107: Fair Use

❑ *Google Books* case

The EU

❑ the Directive on Copyright in the Digital Single Market

❑ Text and data mining exceptions

❑ Article 3

❑ Article 4

The US - Fair Use

❑ *Google Books* case

　　❑ Market-encroaching uses

　　❑ Non-market encroaching uses

The EU - TDM exceptions

❑ Article 3

  ❑ "… exception … for reproductions and extractions made by research organisations and cultural heritage institutions in order to carry out, for the purposes of scientific research, text and data mining of works or other subject matter to which they have lawful access."

The EU - TDM exceptions

❑ Article 4

  ❑ *"The exception or limitation provided for in paragraph 1 shall apply on condition that the use of works and other subject matter referred to in that paragraph has not been expressly reserved by their rightholders in an appropriate manner, such as machine-readable means in the case of content made publicly available online."*

Licensing

❑ Practicalities:

    ❑ Compulsory licensing?

    ❑ A fixed license fee?

Thank you.