POSSIBLE LEGAL RESTRICTIONS ON THE USE OF DATA

Are you training an AI or other system with data? Do you want to know which legal restricitions you can encounter during training? Then be sure to read this brAInfood.

Knowledge Centre Data & Society, 'Possible restrictions on the use of data - brAInfood' (2022), Brussels.

AI FLANDERS
BUILDING OUR DIGITAL FUTURE

DATABASE/-SET

Databases can be protected under copyright and the sui generis database right. In this context a database needs to be arranged in a systematic or methodical way and the contents need to be individually accessible by electronic or other means. This means that unstructured databases/-sets cannot enjoy protection under this regime.

- **Copyright:** if the selection or arrangement of the database content is original, the database structure is protected by copyright. A structure cannot be original if it is dictated by technical considerations or is obvious. Such protection does not extend to the content itself of the database.
- **Sui generis database right:** this protection can be enjoyed if a substantial qualitative or quantitative investment in the obtaining, verification or presentation of the contents can be proven (e.g. if the establishment of the database required intensive data cleaning/labelling/...) .

Vice versa, investments made to generate or create data are irrelevant, ruling out machine-generated (IoT) databases (see also Data Act).

This right only prevents copying and/or publishing of substantial parts of a database by third parties. Vice versa, insubstantial parts may be copied or published, while they are also allowed to consult a database without risking infringement.

Disclaimer: this protection can only be enjoyed by EU citizens and companies.

DATA

Personal data has to be processed in accordance with the **GDPR** and **related national regulation** (e.g. personal image rights or specific rights in relation to employee data).

Network traffic data and location data have to be processed in accordance with **ePrivacy rules** (especially relevant for telecom-organisations).

DATA STRUCTURES/FORMATS

New data structures/formats may be protected under **patent law** if they have a technical effect and fulfil the other patentability requirements.

DATA SOURCE/USER = REGULATED OCCUPATION?

If the data originates from or is used in the context of a regulated occupation (e.g. doctors, accountants or architects) **specific deontological or professional secrecy obligations** may apply restricting the possible (re-)use of data.

DATA = WORKS

If the content of a database/-set consists of **literary or artistic works**, **copyright protection** may apply to those individual works if they are original. Their reproduction or publication requires permission from the copyright holder.

EU Member States are in the process of adopting a **Text-and Data Mining (TDM) exception** which allows the automated analysis of text and data in digital form in order to generate information including e.g. patterns, trends and correlations.

This exception is especially relevant
- for research and/or cultural heritage institutions, and
- to the extent the rightholder has not restricted TDM through contractual or machine-readable means (see publicly available data).

a database/-set

TECHNICAL PROTECTION MECHANISMS/FACTUAL RESTRICTIONS

Data providers may opt to impose **technical protection mechanisms or other factual restrictions** on the use of data. Think about encryption, access restrictions (e.g. passwords/product keys or limited simultaneous use), copy prevention (e.g. download, copy or print blocking, proprietary viewer), rights management software, robots exclusion protocol,...

DATA-AS-A-SERVICE/CONTRACTUAL RESTRICTIONS

As IP protection is often unavailable for data(-sets), data providers may also opt to protect them as **trade secrets**. This means that data providers can prevent the misappropriation of data(-sets) considered to be a trade secret if they took appropriate measures to ensure the secrecy of the data(-set). In practice, this will often result in a **Non-Disclosure Agreement** between the parties, stipulating e.g. access and use modalities, strict confidentiality obligations and possibly ruling out reverse engineering of the data(-set) (which is not prohibited by trade secret law).

More generally, and especially in the context of Data-as-a-Service, data providers will likely require data customers to enter into **contractual licenses**. Generally speaking such contractual terms will include provisions regarding data access and use. Said provisions can specify who, how, for what purpose, for which duration,... data may be accessed and/or used (incl. read-, write- and commit-privileges). Such contract will very likely state what should happen with the data and/or any derived data when the service is terminated (e.g. deletion or restitution).

Data(-sets) may be made available under **open source (OS) licenses**. It should always be ascertained which downstream limitations apply (e.g. attribution, adaptation, commercialization, same OS-license,...) and the mutual compatibility between the applicable OS-licenses.

PUBLICLY AVAILABLE DATA

Even though data may be publicly available (online), data providers may turn them into products and market them under a variety of the restrictions discussed on this **leaflet**.

Moreover, it is important to understand that such data may be subject to e.g. an **online End User License Agreement** (EULA) or **Terms of Use** of a website. Such terms can limit the allowability of TDM (see Data = Works).

SOON: AI ACT - DATA ACT

Further statutory restrictions or obligations regarding the use of data(-sets) may apply in the near future.
- The **AI Act** will very likely impose some sort of data management and data governance requirements in the context of AI applications.
- The **Data Act** will presumably include business-to-government data sharing obligations (in certain circumstances). It will exclude sui generis database protection for machine-generated databases.