Kia kaha! Stay strong!

Data licensing and open science with a particular focus on energy policy making

EERAdata WS5 workshop
Sustainable models for FAIR and open low carbon data

Tuesday 04 October 2022 ▪ 09:45 CEST ▪ Brussels, Belgium

Robbie Morrison
robbie.morrison@posteo.de
Schillerstraße 85, 10627 Berlin, Germany

Release 01 ▪ 03 October 2022
Git information: b12b6cc ▪ 2022-10-03 23:10:25 +0200 ▪ ~/synk/academic/eeradata-2022oct-brussels/eeradata-2022oct-presentation
Generated file: 2022-morrison-eeradata-data-licensing-slidedeck.01.pdf

Abstract

Suitably licensed open data together with relevant free data standards can help improve public policy for the much needed transition to net-zero energy systems. And this is more than simply a matter of providing analysts with access to usable and reusable information, genuine open data and associated standards are necessary conditions for developing a broad and coherent knowledge commons.

This presentation covers the legal context for energy data generally, the role of public licensing, the role of free standards, the need for collective agreement on semantics and metadata in particular, and benefits of community curation. The presentation is limited to non-personal data that may be made public legitimately, thereby avoiding matters of personal and commercial privacy. There is also an emphasis on European Union law.

Data license choice is traversed and, in the absence of further considerations, the Creative Commons CC-BY-4.0 license is recommended for primary data and the Creative Commons CC0-1.0 waiver for metadata and cataloging information. Data standards should be dual licenced with CC-BY-4.0 + MIT — with the latter being a software license that covers potentially patentable material.

▢

Health warning!

In respect of intellectual property held in common, there is little to draw upon in terms of:

- supportive legislation
- case law
- official interest
- academic analysis

Preamble

Some background

- 1990 : began campaigning on global warming in a personal capacity
- 1995 : modeling national energy systems: multi-sector, high-resolution, contiguous time
- 2003 : added the GPL-2.0 license to *deeco* and attempted to build an online community
- 2016 : joined the Open Energy Modelling Initiative (openmod)
- 2017 : joined the Free Software Foundation Europe (FSFE) Legal Network (LN)
- 2017 : began lobbying for genuinely open data of public interest — a clear Achilles heel

- have coordinated four written submissions on data law reform in Europe: [1]
  - revised Public Sector Information (PSI) Directive, now Open Data Directive 2019/1024
  - proposed Data Act

[1] Listed at rear of the presentation (slide 59)

Context

Quantitative information of public interest

This presentation covers **non-personal data** that is of **public interest** and can be or has been **legitimately published** — hence:

Necessarily excluded:

- **personal data** — personally identifiable information (PII)
- **trade secrets** — which can include "compilations of information"
- **shared data** — consortium data and brokered data

Included:

- material under **statutory reporting** — noting that most mandates seek to assist system security or remedy market failure rather than advance sustainability
- collaborative projects leveraging **citizen science**:
  - OpenStreetMap — ODbL-1.0 license
  - Wikidata — CC0-1.0 license

Open data

"Open definition" — Open Knowledge Foundation short-form

Knowledge is open if anyone is free to access, use, modify, and share it — subject, at most, to measures that preserve provenance and openness

Noting: [1]

- "share [then] use [and] modify" supports "reuse" and "reutilization"
- "preserve provenance" implies "attribute" and the Creative Commons **BY** trait
- "preserve openness" implies "sharealike" and the Creative Commons **SA** trait [2]

Long-form open definition available at:

- Open Knowledge Foundation (no date). Open Definition 2.1 — Defining open in open data, open content and open knowledge. *Open Knowledge Foundation (OKF)*. Cambridge, UK.

[1] See also the "open data" category on the openmod forum (circa 60 threads): https://forum.openmod.org/c/open-data

[2] The associated CC-BY-SA-4.0 license is not generally recommended for data. Indeed no sharealike licenses are.

Legal scope

Most remarks here pertain to **European Union law** in aggregate — noting that EU directives are necessarily transposed into national law and the resulting legislation is not necessarily harmonized

Broad areas of law:

- law on **intellectual property rights** (IPR) — the focus here
- law of **contracts** – although not in the United Kingdom [1]
- law on **business wrongs** (torts) — not discussed further here
- law covering **injunctions against intermediaries** — those providing portals [2]

More specific areas:

- legislation mandating **statutory reporting** — often for system security and market failure
- legislation covering use of **public sector information**

---

[1] United Kingdom contract law requires a reciprocal consideration (such as a payment) for a contract to form and therefore explicit contact between the licensee and licensor — the other branches of law apply in the UK notwithstanding

[2] No United States §230 "safe harbor" provision in Europe

Not traversed

Not covered in this presentation:

- outputs generated by **artificial intelligence** methods are expressly excluded unless the ingested material is suitably open licensed or clearly unencumbered [1]
- activities deemed **lawfully permitted** as **exceptions** under copyright and allied rights law — for instance **scientific research** and **data mining** under German legislation — produce data that must necessarily remain corralled (and this doubtless impedes the migration to open science and contestable policy formation)
- the role of **research funding policy** — but arguably somewhere where significant progress can be made

[1] Refer to: Hugenholtz, P Bernt and João Pedro Quintais (1 October 2021). "Copyright and artificial creation: does EU copyright law protect AI-assisted output?". *International Review of Intellectual Property and Competition Law*. **52** (9): 1190–1216. ISSN 2195-0237. doi:10.1007/s40319-021-01115-0. Open access.

Data standards

Data standards are often neglected in the debate over genuinely open data

These cover three main areas:

- **data semantics** — including glossaries, reference architectures, ontologies
- **technical specifications** — including technical interchange
- **legal interoperability** — more later

Free standards and open standards differ:

- **free standards** are published under CC-BY-4.0 + MIT together [1]
- so-called **open standards** are often proprietary and offered under **FRAND** (fair, reasonable, and non-discriminatory) terms but otherwise protected as intellectual property → say with a licensing fee of €2000

FRAND does not work for inclusive open modeling communities

[1] The MIT license provides a patent grant which CC-BY-4.0 alone does not

Standards setting more generally

Indeed, broader tensions exist between **standards setting organizations** (SSO) and open data maintainers. Boehm and Eisape (2021) review similar potential conflicts in the context of open source software [1]

The general issue of free and usable standards needs more attention

[1] Boehm, Mirko and Davis Eisape (28 June 2021). "Standard setting organizations and open source communities: partners or competitors?". *First Monday*. **26** (7): 10806. ISSN 1396-0466. doi:10.5210/fm.v26i7.10806. Public Domain.

Some research trends

Generalizing somewhat, but **energy systems researchers** in Europe:

- often work in increasingly **legally risk averse** environments
- are adopting **open science** doctrines, including **strict reproducibility**
- increasingly develop and use **custom software**
- are recognizing the benefits of **collaborative development** for both software and data
- are **widely reliant** on what the European Commission describes as "privately-held information [of] public interest"
- those developing **open source frameworks** are desperate for **genuinely open and semantically consistent data**
- those working with **open models** are progressively venturing into the previously **closed world** of **public policy analysis**
- some modeling teams are now working with researchers from the **global south**

Representative community projects

Here are some dedicated community projects in the energy sector centered on data management and increasingly looking toward linked open data (LOD):

Europe:

- Open Energy Platform (OEP) — also strong focus on semantics and technical standards
- Open Power System Data (OPSD) — pulls from the ENTSO-E Transparency Platform [1]

United States:

- PowerGenome — pulls from the Public Utility Data Liberation (PUDL) project and US EIA

See also Wikipedia on Open energy system databases

[1] The portal indicates that users wishing to republish original or modified datasets should seek permission from the "primary data owner" — a term defined in regulation 543/2013 without reference to which intellectual property rights that might apply as follows at §2.23: " 'primary owner of the data' means the entity which creates the data"

Ultimate goal is a knowledge commons

This presentation assumes the objective is to create a **knowledge commons** comprising: [1]

- fully usable and re-usable data
- community curation of **canonical** data
- consensus semantics
- underpinning technical standards that are free
- necessarily reliant on **distributed architectures** and **linked open data** (LOD) methods

[1] Hoyer-Klick, Carsten, Johannes Frey, Ulrich Frey, Hedda Gardian, Anastasis Giannousakis, Jan Göpfert, Tobias Hecking, Christian Hofmann, Sophie Jentzsch, Kevin Knosala, Leander Kotzur, Stefan Kronshage, Patrick Kuckertz, Christoph Muschner, Michaja Pehl, Vera Sehn, and Detlef Stolten (28 October 2021). *Implementing FAIR through a distributed data infrastructure*. Germany: DLR *et al.* Parallel session presentation to EMP-E 2021 online conference, 28 October 2021, 14:00–15:30 CEST.

Intellectual property rights

IPR that does not require examination, grant, and payment (unlike trademarks and patents)

Copyright

- a legally sanctioned time-limited private monopoly right
- general copyright — was developed for literary works and extended to other media
- special case legislation too — the legal protection of computer programs for example
- so-called moral rights apply in Europe — such as the right to be associated with the work

Database protection

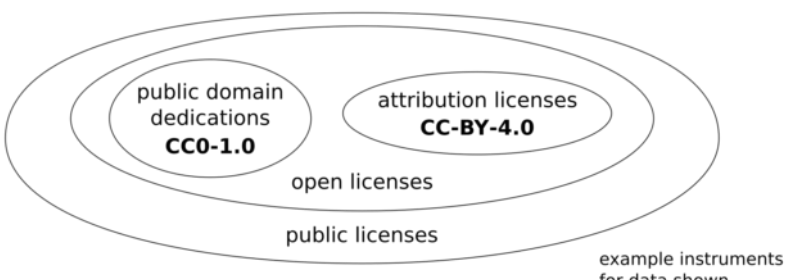
- Database Directive 96/9/EC introduced in 1996
- covers the European Economic Area (EEA) and currently post-Brexit United Kingdom
- intended to safeguard a fledgling computer database industry
- set within wider aspirations to expand the market for information products and services
- no equivalent legislation in the United States, despite several attempts
- widely disliked today

Intellectual property held in common

Virtually no legislative support whatsoever for intellectual property held in common

This led to third-party **public licenses** to provide the necessary permissions and restrictions:

- first for software : GNU GPL family in 1989
- later for data : the Creative Commons family from version 4.0 being data-capable



Public sector information in Europe

Public sector information (PSI) is covered by the **Open Data Directive** 2019/1024 (ODD)

- ODD built on earlier legislation
- intended to enable better use of information generated by "public sector bodies"
- public sector bodies can no longer claim 96/9/EC database protection (§1.6)
- **research data** from universities now covered under the rubric of "open access" [1]

Identifying a **public sector body** is more difficult and restrictive than one might imagine:

- ENTSO-E — the transmission system operators umbrella organization — is not [2]

[1] Under ODD recital 27, the term "open access" is "understood as" free-of-charge online access "without restrictions on use" beyond optionally the need to acknowledge — also a considerably more liberal treatment of the term than often found

[2] ENTSO-E was established under regulation 714/2009 concerning conditions for access to the network for cross-border exchanges in electricity / legal status of ENTSO-E as not being a public sector body confirmed to me by letter

Open Data Directive 2019/1024 and data re-use ***

Definition §2.11

"'re-use' means the **use** by persons or legal entities of documents held by . . . "

This remapping of "re-use" → "use" is doubtless **problematic**:

- "use" is a well understood and well delineated concept within intellectual property law
- the right granted to "use" **does not provide the general right to copy and republish** in original or modified form — those particular activities lie well outside documented exceptions under European law or affirmative defenses under fair use in the United States

My interpretation therefore:

- the concept of "re-use" in the ODD is **restricted to the first hop** from the PSI provider
- **no rights are granted to copy and republish that material** in original or modified form
- entirely counter to established norms for "open data" (see Open Knowledge Foundation)

Copyright verbs

A **copyright verb** is an established legal concept:

- used capture what a **recipient** may and may not do

Examples of copyright verbs in relation to **numerical data** taken from legislation, court judgments, and open licenses (others sets of verbs and definitions apply to source code, computer programs more generally, and other forms of media):

Copyright verb	Comment
copy	reading data into computer RAM storage is an act of copying
distribute	
extract	related to 96/9/EC databases
modify	
prepare derivative works	derivative works retain any prior copyright
publish	
redistribute	
reuse | re-use	mapped to mere "use" under definition §2.11 of the Open Data Directive
reutilize	related to 96/9/EC databases
share	
share adapted material	
use	local numerical processing but not redistribution

The wider policy setting

Clear tensions between:

- the current drive to create a European **data market** for data products and services [1]
- growing interest in creating a European **digital commons** of sorts
- some desire to create an ambitious **digital twin** of the European energy system [2,3]

France, who holds the current EU presidency from 2022, recently indicated a desire to establish a "digital commons" (emphasis added) [4]

> *The digital commons utilize an **open approach** and are based on the **collective control** and use of data and technological infrastructure.*

---

[1] The proposed Digital Markets Act, scheduled for 2023, represents one current iteration

[2] Bertuzzi, Luca (14 September 2022). Leak: EU prepares 'action plan' to digitalise the energy system. *www.euractiv.com*.

[3] Overarching projects like this can easily divert attention from fixing existing problems — of which there are many

[4] French Embassy (7 February 2022). France calls for a European initiative for digital commons. *France in the UK*. London, UK.

Peer production and public policy formation

There is no reason why open analysis cannot contribute to public policy formation — this has begun but there is little to no official recognition

Indeed the European Commission:

- identifies **commons-based peer production** as a megatrend [1]
- but completely fails to embrace the concept for public policy formation

[1] Warnke, Philine, Kerstin Cuhls, Ulrich Schmoch, Lea Daniel, Liviu Andreescu, Bianca Dragomir, Radu Gheorghiu, Catalina Baboschi, Adrian Curaj, Marjukka Parkkinen, and Osmo Kuusi (3 December 2019). *100 radical innovation breakthroughs for the future — KI-04-19-053-EN-N*. Luxembourg: Publications Office of the European Union. ISBN 978-92-79-99139-4. doi:10.2777/24537. Pages 259–261.

Legal issues and examples

Data object for analysis

To assist with analysis, one type of **transactional object** is mooted for discussion:

Plain text tabular dataset

- a collection of observed atomic items: namely numbers or facts
- an associated schema — implied or explicit
- and perhaps accompanying metadata — ideally employing a standard vocabulary [1]
- entirely **passive** and **human readable** using a text editor

- hence consider a one file OKF **frictionless data package** comprising: [2]
  - plain text CSV tabular data
  - YAML or JSON-specified table schema and information on CSV dialect
  - accompanying metadata, also notated in YAML or JSON
  - possibly compressed using the gzip utility

[1] Perhaps the DCAT data catalog vocabulary and/or the Dublin Core Metadata Element Set (DCMES) as appropriate

[2] Frictionless data website: https://specs.frictionlessdata.io

Copyright for collections of atomic data

***

Consider our entirely passive plain text **tabular dataset** from a copyright perspective:

- it classes as a **collection** (or compilation) of non-copyrightable elements: the **atomic data**
- may have multiple **creators** whose contributions cannot be distinguished: **joint authorship**
- may have been modified or combined with other datasets: thus a **derivative work**

To attract copyright in its own right, normally:

- real **humans** must have generated the contents
- some minimum **threshold of originality** must have been reached or exceeded
- the above doctrine naturally excludes **trivial works**
- **no protection** when originality is insufficient — and that includes public sector information

The threshold of originality varies by legal jurisdiction and evolves with case law. Under Germany law, copyright attaches to a collection if and only if (emphasis added) (UrhG §4.1):

> *the "**selection and arrangement**" of the elements is sufficiently creative*

Copyright revisited — and our tabular dataset

Copyright in a collection as per our minimal tabular dataset is **unlikely** for energy sector data — but we just don't **know** for sure

- a tabular dataset would doubtless count as a collection under copyright law
- whether it meets the threshold of originality is another matter

- some published data is from SCADA systems or market clearance and dispatch algorithms
- other published data comprises entirely routine lists of information such as physical assets
- I would guess most examples do not reach the threshold of originality for protection

- unrelated creative material can be payloaded in to ensure copyright protection [1]

[1] Osborn, Lucas S (2017). "The limits of creativity in copyright: digital manufacturing files and lockout codes". *Texas A&M Journal of Property Law*. **4**: 25. The practice of including creative material to trip copyright is known as adding "lockout codes".

Database Directive 96/9/EC — overall character

General:

- also known as a "related right" (regarding copyright) or "sui generis" (one of a kind) right
- directive adopted in 1996 and subsequently written into national legislation
- the legal protection covers the database but not its contents
- the definition of a database is wide (emphasis added):

  a collection of independent works, data or other materials arranged in a ***systematic or methodical way*** *and* ***individually accessible*** *by electronic or other means*

Noting that:

- any computer program used to generate the database is excluded from this protection
- an analog object, such as a mass-printed topographical map, may class as a database [1]

And also:

- material served under statutory reporting is **not** expressly excluded by law

[1] Schweizer, Mark (5 November 2015). C-490/14 — Verlag Esterbauer: Get off my map!. *The IPKat*. London, United Kingdom.

Database Directive 96/9/EC — criteria for protection

Two-step requirement:

- for protection — the direct **investment** must be substantial
- for infringement — the **extraction** must be substantial

Noting that:

- the investment criteria is restricted to the database and excludes its contents [1]
- recent case law essentially limits protection to providers facing commercial risk [2,3]
- databases provided by public sector bodies now expressly excluded by the ODD

Case law on these and related matters is slowly emerging — but most is commercial, not public interest, in nature

---

[1] European Court of Justice (9 November 2004). *Judgment of the Court (Grand Chamber) of 9 November 2004 — Case C-203/02 — ECLI:EU:C:2004:695*. Kirchberg, Luxembourg: European Court of Justice (ECJ). Judgment counter to published opinion of Advocate General Stix-Hackl. The so-called BHB case.

[2] Giannopoulou (2018:106) (reproduced as ancillary material on slide 63)

[3] CJEU (3 June 2021). *Judgment of the Court on 'CV-Online Latvia' SIA v 'Melons' SIA, case C-762/19, document ECLI:EU:C:2021:434*. Luxembourg City, Luxembourg: Court of Justice of the European Union (CJEU). 6 pages.

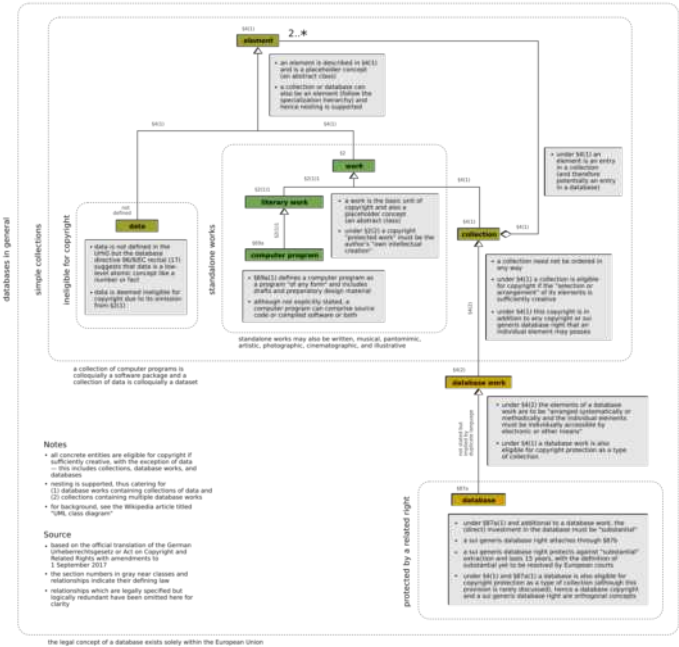
96/9/EC database protection — interpretation

96/9/EC database protection is our **Achilles heel** — users simply cannot know where the legal thresholds for protection for individual portals might lie

- indeed a tabular dataset would comply with the definition of a 96/9/EC database
- the focus here is mostly directed toward official sites
- it is not possible for users to estimate "substantial" extraction
- the scope of a set of databases can be strategically manipulated to maximize protection [1]

[1] Davidson, Mark J (January 2008). *The legal protection of databases*. Cambridge, United Kingdom: Cambridge University Press. ISBN 978-0-521-04945-0. Paperback edition.

Definitions from the German copyright act (Urhg)
UML class diagram



Problematic examples 1

ENTSO-E **Transparency Platform**

- mandated under regulation 543/2013 but legally encumbered information nonetheless
- openmod people have pushed for change over several years but no real movement [1]
- "public sector body" status of ENTSO-E has not been publicly clarified by the organization

Open Power System Data (OPSD) portal

- community site: https://open-power-system-data.org
- draws from ENTSO-E Transparency Platform
- site carries caveat about need to seek permission for re-use from the "primary data owner"

[1] Those involved include LH, ES, IS, TB, myself

Problematic examples 2

Organization	Portal or source	Local
European Commission Joint Research Centre	Energy and Industry Geography Lab	post 6
IIASA	IAMC 1.5°C Scenario Explorer NGFS Scenario Explorer	post 11
International Energy Agency	recent IEA reports	post 5
UNFCCC	website	post 2
World Bank Group	energydata.info portal	post 4
UNEP World Conservation Monitoring Centre	World Database of Protected Areas	post 3

Unclear, ambiguous, and/or contradictory licensing by public bodies [1]

With reference to the **IIASA databases** cited above, IIASA recently reviewed their licensing policy. IIASA hold the view that these databases attract 96/9/EC protection and that the now non-open modified CC-BY-4.0 licensing applied does not inhibit research. This example highlights the tension between wishing to protect against forks and serving genuinely open data to support a knowledge commons.

[1] Sourced from: https://forum.openmod.org/t/3102

Problematic examples 3

Lack of open provision of **day-ahead power prices** per bidding zone — issues include:[1]

- Transparency Regulation 543/2013 foresees this information being published on the ENTSO-E Transparency Platform under statutory reporting
- the various energy exchanges (EXs) and nominated electricity market operators (NEMOs) generate this information in the context of coupled markets and price discovery
- this information is calculated using regulated processes and is of public interest
- while the NEMOs claim this information is proprietary and demand payment for access
- unresolved after 3 years

Also the website of the **EEX** (European Energy Exchange) is

- technically encumbered — copy/pasting intentionally blocked, maybe via CSS or JavaScript
- complaints to **ACER** (European Union Agency for the Cooperation of Energy Regulators) determined that this practice is compliant with statute

[1] Schmid, Eva, Ingmar Schlecht, Tomas Šumskas, and Florence Melchior (October 2019). *Why do we need an open data licence on the ENTSO-E Transparency Platform? An FAQ*.

Recap — main legal concepts

Copyright for a **collection** applies when (from Germany copyright act):

> the "**selection and arrangement**" of the elements is sufficiently creative

The **threshold of originality** varies by jurisdiction:

> sweat-of-the-brow (UK) → non-trivial creativity → individual character required (AT)

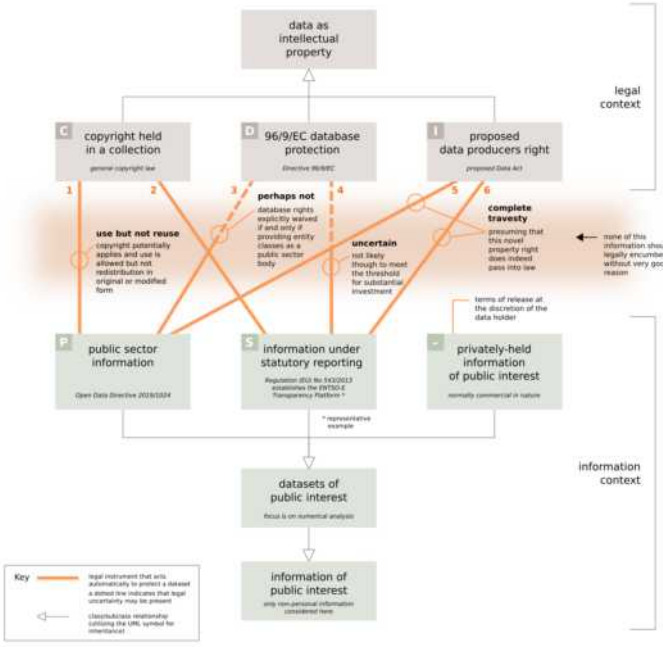
The definition of a **96/9/EC database** as follows:

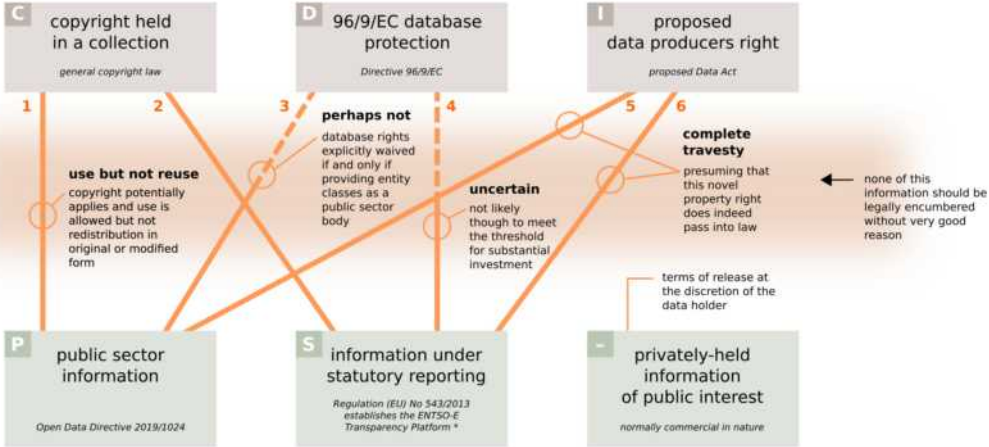
> "a collection of independent works, data or other materials arranged in a **systematic or methodical way** and **individually accessible** by electronic or other means"

The thresholds for **substantial investment** and **substantial extraction** remain unclear — although case law suggests exposure to commercial risk may be required

Both types of property right attach automatically and must be **explicitly removed** through either public licensing or official notice

Information sources meet intellectual property rights
The Data Producers Right has been dropped from the latest draft of the proposed Data Act

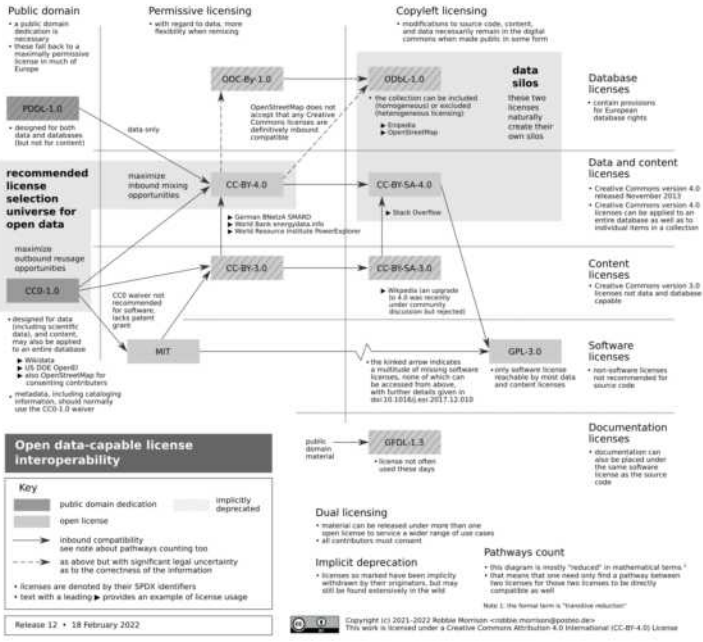


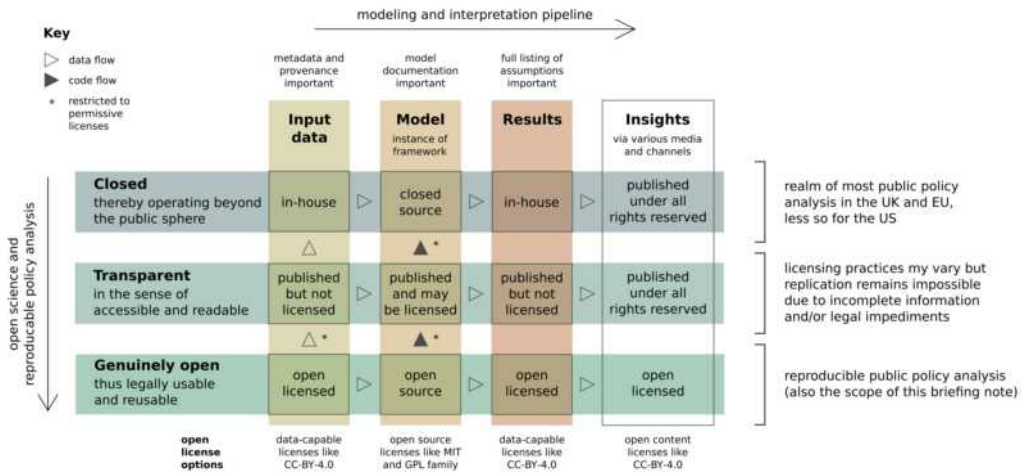

C
copyright held in a collection
general copyright law
D
96/9/EC database protection
Directive 96/9/EC
I
proposed data producers right
proposed Data Act
1
2
3
4
5
6
use but not reuse
copyright potentially applies and use is allowed but not redistribution in original or modified form
perhaps not
database rights explicitly waived if and only if providing entity classes as a public sector body
uncertain
not likely though to meet the threshold for substantial investment
complete travesty
presuming that this novel property right does indeed pass into law
none of this information should be legally encumbered without very good reason
terms of release at the discretion of the data holder
P
public sector information
Open Data Directive 2019/1024
S
information under statutory reporting
Regulation (EU) No 543/2013 establishes the ENTSO-E Transparency Platform *
-
privately-held information of public interest
normally commercial in nature

Open licenses for data

Open data license compatibility chart

Diagram is provisional





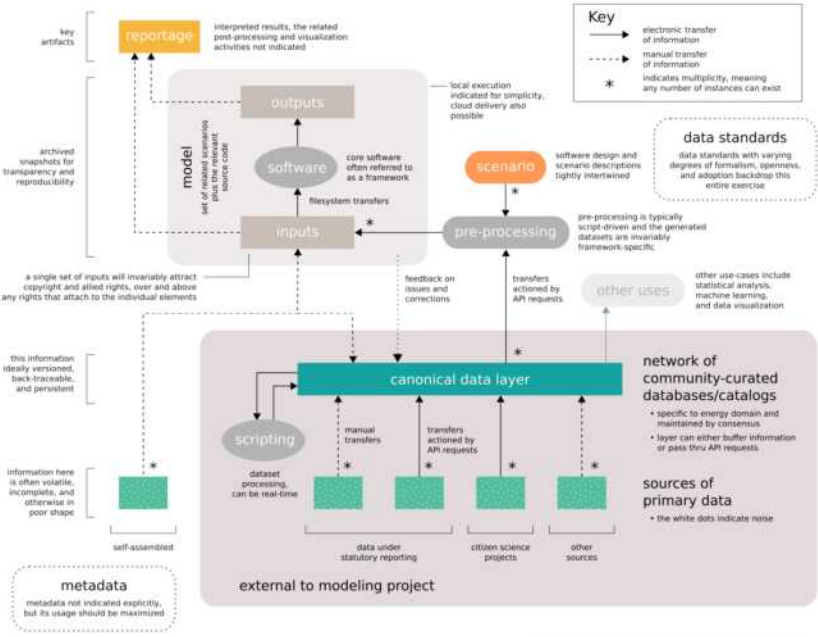
Open source models and genuinely open data naturally pair

Or more forcefully put, a shift to open models demands that data be under CC-BY-4.0 licensing

Knowledge commons

Energy system data flows given a canonical data layer

Towards a knowledge commons, LOD architectures not indicated



Stepping back

From a perspective of energy systems analysis ***

Legislative issues:

- the definition of "re-use" concerning **public sector information** seems entirely deficient
- a "**public sector body**" can be hard to identify when non-traditional
- almost all current **statutory reporting** is potentially legally encumbered [1]

Technical issues

- most energy sector datasets unlikely to be sufficiently original to attract copyright ?
- databases established by statute unlikely to retain 96/9/EC protection if brought to court ?

Looking forward

- the push by France for a European digital commons may work in our favor if accompanied by suitable licensing ?

[1] Exceptions include the BNetzA SMARD portal under CC-BY-4.0 and the French RTE portal under CC-BY-4.0

Reality check

Labastida and Margoni (2020:205) opine: [1, 2]

> *It should be clarified that in many instances there will be no copyright or related rights on data.*

---

[1] Labastida, Ignasi and Thomas Margoni (1 January 2020). "Licensing FAIR data for reuse". *Data Intelligence*. **2** (1-2): 199–207. ISSN 2641-435X. doi:10.1162/dint_a_00042. CC-BY-4.0 license.

[2] The terms allied rights"" and "related rights" refers to 96/9/EC database protection

The tragedy of the anticommons

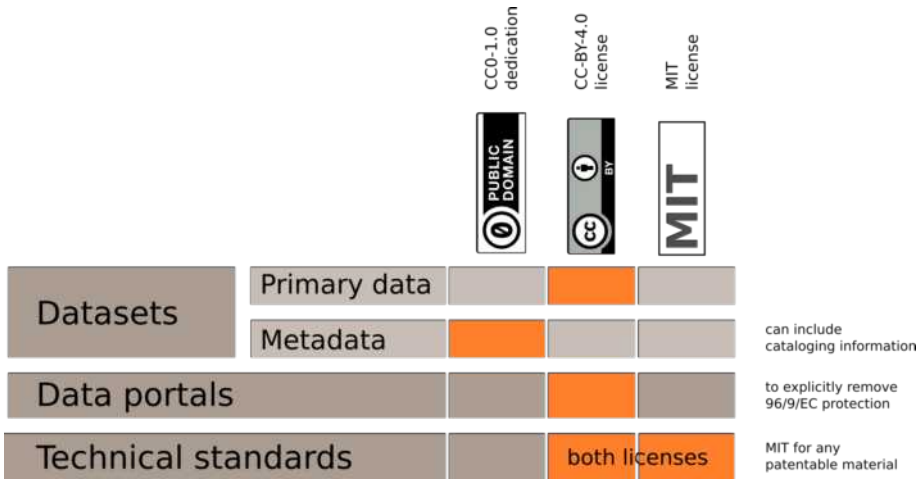
The **tragedy of the anticommons** is a type of coordination breakdown, in which a commons does not emerge, even when the general access to resources or infrastructure would be a social good (*source*: Wikipedia)

Indeed, the idea of a genuine **knowledge commons** to serve public interest analysis does not seem to come easily to legislators

Rather, supplying an emerging **data market** with potentially non-reusable public sector information would appear to be uppermost in mind

Solutions

Explicit open licensing on datasets / portals / data standards



Explicit open licensing — rationale

In **most cases**, open licenses do not provide users with **permissions** — because the underlying material is not intrinsically protected — but they do provide users with **legal certainty**

My wish list for the European Commission

Looking to the horizon, the **European Commission** should:

- help repair the Open Data Directive 2019/1024 — particularly the definition for "re-use"
- expand the rationale for statutory reporting to include the rapid transition to a more sustainable society — and make that reporting genuinely accessible and re-usable by law
- develop legislative support for open intellectual property — covering both code and data — thereby removing much of the need for the current raft of both community and official third-party public licenses offering second best solutions
- buy out the copyrights for key data standards — and make them available as free standards
- analyze community and official public data license compatibilities (as per earlier digraph diagram) — a task requiring painstaking legal scrutiny

"Data leakage" to United States servers

If suitable open licensing is not forthcoming, published PSI datasets will doubtless "leak" to public-interest data portals located in the United States and be granted CC-BY-4.0 or similar:

- Mireille van Eechoud (2021:378) covers this scenario and opines that the proposed Data Governance Act lacks clarity in respect of public sector information [1]
- such "data leakage" already occurs — the US-based World Resources Institute (WRI) republishes datasets drawn from the ENTSO-E Transparency Platform and serves them under CC-BY-4.0 licensing

In the absence of so-called "adequacy requirements", the location of the server determines the intellectual property law that applies. And United States law in respect of datasets and databases is comparatively lax and neither are likely to attract IPR protection in this context [2]

---

[1] At the time of writing, that Data Governance Act has yet to be formally approved

[2] US Copyright Office (November 2017). *The Compendium of US Copyright Office Practices — Third edition: Chapter 700*. US Government. Refer §727 and specifically §727.1

Data reusability is a legal swamp that the EC can help fix

Reference matter

Abbreviations

ACER	Agency for the Cooperation of Energy Regulators
CSV	comma-separated values
DCMES	Dublin Core Metadata Element Set
DG	European Union Directorate-General
EC	European Commission
ECJ	European Court of Justice
EEA	European economic area
EEX	European Energy Exchange
EIA	US Energy Information Administration
FRAND	fair, reasonable, and non-discriminatory
IEA	International Energy Agency
IPR	intellectual property right
JSON	JavaScript Object Notation
LOD	linked open data
ODD	Open Data Directive 2019/1024
OEP	Open Energy Platform
OKF	Open Knowledge Foundation
openmod	Open Energy Modelling Initiative
OPSD	Open Power System Data
PSI	public sector information
PUDL	Public Utility Data Liberation project
REMIT	Regulation on Wholesale Energy Market Integrity and Transparency
SCADA	supervisory control and data acquisition
SPDX	Software Package Data Exchange (provides unique public license identifiers)
UrhG	Urheberrechtsgesetz (German copyright act)
YAML	yet another markup language (a human-readable data-serialization language)

Selected legislation

- European Parliament and European Council (27 March 1996). "Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases". *Official Journal of the European Union*. **L 77**: 20–28. Established so-called *sui generis* database right.
- European Commission (14 August 2009). "Regulation (EC) No 714/2009 of the European Parliament and of the Council of 13 July 2009 on conditions for access to the network for cross-border exchanges in electricity and repealing Regulation (EC) No 1228/2003 (text with EEA relevance)". *Official Journal of the European Union*. **L 211**: 15–35. Established ENTSO-E.
- European Commission (8 December 2011). "Regulation (EU) No 1227/2011 of the European Parliament and of the Council of 25 October 2011 on wholesale energy market integrity and transparency (text with EEA relevance)". *Official Journal of the European Union*. **L 326**: 1–16. Established Regulation on Wholesale Energy Market Integrity and Transparency (REMIT).
- European Commission (15 June 2013). "Commission Regulation (EU) No 543/2013 of 14 June 2013 on submission and publication of data in electricity markets and amending Annex I to Regulation (EC) No 714/2009 of the European Parliament and of the Council (text with EEA relevance)". *Official Journal of the European Union*. **L 163**: 1–12. Established the ENTSO-E Transparency Platform.
- European Commission (26 June 2019). "Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information (recast)". *Official Journal of the European Union*. **L 172**: 56–83. Replaced an earlier public sector information directive.

Legislation in progress

Proposed Data Act:

- European Commission (23 February 2022). *Data Act: Proposal for a Regulation on harmonised rules on fair access to and use of data (text with EEA significance) — COM(2022) 68 final*. Brussels, Belgium: European Commission.
- European Commission (23 February 2022). *Commission staff working document on common European data spaces — SWD(2022) 45 final*. Brussels, Belgium: European Commission. Sole focus is creating venues to facilitate information transactions.

Some readings / 1

- Anon (24 January 2020). *B2 — Analytical report on EU law applicable to sharing of non-personal data — V2.0*. Capgemini Invent, Fraunhofer FOCUS, Timelex, Support Centre for Data Sharing. Report for DG Connect (DG = European Union Directorate-General).
- Bimesdörfe, Kathrin (editor) (February 2019). *Datenlizenzen für Open Government Data: Rechtliches Kurzgutachten: Handreichung zu den Nutzungsrechteregelungen gebräuchlicher Open Data Lizenzen und Empfehlungen für ihren Einsatz* [*Data licenses for Open Government Data: Legal brief: Guidance on the usage rights of common open data licenses and recommendations for their use*] (in German). Düsseldorf, Germany: Ministerium für Wirtschaft, Innovation, Digitalisierung und Energie des Landes Nordrhein-Westfalen.
- Chestek, Pamela S (2017). "A theory of joint authorship for free and open source software projects". *Colorado Technology Law Journal*. **16**: 285–326. Open access. The arguments apply equally to datasets with multiple contributors.
- Davidson, Mark J (January 2008). *The legal protection of databases*. Cambridge, United Kingdom: Cambridge University Press. ISBN 978-0-521-04945-0. Paperback edition.
- van Eechoud, Mireille (1 April 2021). "A serpent eating its tail: the Database Directive meets the Open Data Directive". *International Review of Intellectual Property and Competition Law*. **52** (4): 375–378. ISSN 2195-0237. doi:10.1007/s40319-021-01049-7. Editorial. Open access.

Some readings / 2

- Giannopoulou, Alexandra (2018). *Chapter 6: Understanding open data regulation: an analysis of the licensing landscape*. In Bastiaan van Loenen, Glenn Vancauwenberghe, and Joep Crompvoets (editors) (2018). *Open data exposed*. The Hague, the Netherlands: TMC Asser Press. Pages 101–125. ISBN 978-94-6265-261-3. doi:10.1007/978-94-6265-261-3_6.
- Hirth, Lion (1 January 2020). "Open data for electricity modeling: legal aspects". *Energy Strategy Reviews*. **27**: 100433. ISSN 2211-467X. doi:10.1016/j.esr.2019.100433. Open access.
- Hoyer-Klick, Carsten, Johannes Frey, Ulrich Frey, Hedda Gardian, Anastasis Giannousakis, Jan Göpfert, Tobias Hecking, Christian Hofmann, Sophie Jentzsch, Kevin Knosala, Leander Kotzur, Stefan Kronshage, Patrick Kuckertz, Christoph Muschner, Michaja Pehl, Vera Sehn, and Detlef Stolten (28 October 2021). *Implementing FAIR through a distributed data infrastructure*. Germany: DLR *et al.* Parallel session presentation to EMP-E 2021 online conference, 28 October 2021, 14:00–15:30 CEST.
- Husovec, Martin (November 2017). *Injunctions against intermediaries in the European Union: accountable but not liable*. Cambridge, United Kingdom: Cambridge University Press. ISBN 978-1-108-41506-4. doi:10.1017/9781108227421.
- Kelly, Jack (14 February 2022). *Data, data, every where; nor any drop to drink*. doi:10.5281/zenodo.6079270. Informal position paper. CC-BY-4.0 license.

Some readings / 3

- Jaeger, Till (24 July 2017). *Legal aspects of European electricity data — Legal opinion*. Berlin, Germany: JBB Rechtsanwälte.
- Labastida, Ignasi and Thomas Margoni (1 January 2020). "Licensing FAIR data for reuse". *Data Intelligence*. **2** (1-2): 199–207. ISSN 2641-435X. doi:10.1162/dint_a_00042. CC-BY-4.0 license.
- van Loenen, Bastiaan, Glenn Vancauwenberghe, and Joep Crompvoets (editors) (2018). *Open data exposed*. The Hague, the Netherlands: TMC Asser Press. ISBN 978-94-6265-261-3. doi:10.1007/978-94-6265-261-3.
- Morrison, Robbie (6 February 2022). *Which open data license? — Release 06*. doi:10.5281/zenodo.5987672. 14 pages. CC-BY-4.0 license.
- Stepanov, Ivan (2 January 2020). "Introducing a property right over data in the EU: the data producer's right — an evaluation". *International Review of Law, Computers and Technology*. **34** (1): 65–86. ISSN 1360-0869. doi:10.1080/13600869.2019.1631621. Open access.

Community submissions

For reference purposes, these are submissions made by the open energy modeling community to the European Commission on proposed legislation:

- Morrison, Robbie, Tom Brown, and Matteo De Felice (10 December 2017). *Submission on the re-use of public sector information: with an emphasis on energy system datasets — Release 09*. Berlin, Germany. Creative Commons CC BY 4.0 license. 23 pages. 39 submitters.
- Morrison, Robbie and Lion Hirth (19 February 2018). *Follow-up to the public consultation on the re-use of public sector information: issues surrounding sui generis database rights — Release 03*. Berlin, Germany. Creative Commons CC-BY-4.0 license. Requested input. 2 pages.
- Morrison, Robbie (25 June 2021). *Submission on a proposed Data Act for the European Union from the perspective of energy system analysis — Release 07*. doi:10.5281/zenodo.5032198. Berlin, Germany. Creative Commons CC-BY-4.0 license. 19 pages. 16 submitters.
- Morrison, Robbie (3 September 2021). *Submission on a proposed Data Act for the European Union from the perspective of energy system analysis / 2 — Release 02*. doi:10.5281/zenodo.5471077. Berlin, Germany. Creative Commons CC-BY-4.0 license. 7 pages. Sole submitter.

Thanks for your attention

Ancillary material

European Union definition for open data

From the open data directive as a recital as opposed to definition §2.11 in the main body:

Open data directive 2019/1024 — Recital 16 [1]

"open data as a concept is generally understood to denote data in an open format that can be freely used, re-used and shared by anyone for any purpose"

[1] European Commission (26 June 2019). "Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information (recast)". *Official Journal of the European Union*. **L 172**: 56–83.

Giannopoulou (2018:106)

The Database Directive does not clearly indicate the exclusion of public databases that fall under the PSI Directive from qualifying for the sui generis protection. In principle, since public sector databases are not excluded, branches of state power can benefit from the sui generis right protection when they fulfill the conditions.[36] Absent an ECJ decision, however, courts from some Member States have ruled against the possibility of public bodies asserting sui generis database rights. Namely, courts in Italy and Germany have held that even if public sector databases qualify for the protection, they should be exempt from it.[37] The highest administrative court in Amsterdam has held that the City of Amsterdam cannot hold sui generis rights on a database even if it has made a substantial investment towards its creation because the has not borne the risk for the investment in question.[38] Thus, it cannot impose limitations or charges in the reuse of that database. Finally, French law has been amended [39] to clarify that public bodies cannot invoke a sui generis right in order to refuse the reuse of their data.

Giannopoulou, Alexandra (2018). *Chapter 6: Understanding open data regulation: an analysis of the licensing landscape*. In Bastiaan van Loenen, Glenn Vancauwenberghe, and Joep Crompvoets (editors) (2018). *Open data exposed*. The Hague, the Netherlands: TMC Asser Press. Pages 101–125. ISBN 978-94-6265-261-3. doi:10.1007/978-94-6265-261-3_6. The analysis above predates the open data directive 2019/1024.

Free and open standards in general

What are the legal requirements on unencumbered standards:

- US Supreme Court 2021 ruling on reimplemented public APIs that “fair use” applies irrespective but silent on whether copyright attached

Google LLC v Oracle America, Inc. Docket no. 18-956. Decided 5 April 2021.

Photographs

Swamp in Spandau

- description: Teufelsbruch wetland, Spandauer Forst, Spandau, Berlin, Germany
- timestamp: 2021-02-28 10:45:51+00:00
- lat/lon: +52.5796 +013.1994
- conditions: Sony ILCE-6600 ▪ 15mm (35mm equivalent) ▪ ISO:100 1/60 f4.0
- photographer: Robbie Morrison
- image: STR05750.JPG

Garlic mustard close-up

- description: Garlic mustard (Alliaria petiolata) [Knoblauchrauke], Brandenburg, Germany
- timestamp: 2021-05-29 08:24:22+00:00
- conditions: Sony ILCE-6600 ▪ 45mm (35mm equivalent) ▪ ISO:125 1/60 f4.0
- photographer: Robbie Morrison
- image: STR08402.JPG