

Understanding law and licensing collecting, using and publishing data

Francis Davey

Open Data Institute
65 Clifton Street
London
EC2A 4JE
contact@francisdavey.co.uk

Overview

Who is this directed at?

- ▶ Consumers
 - ▶ Collecting
 - ▶ Using
- ▶ Producers
 - ▶ Publishing

Topics

Only two topics today.

- ▶ Property rights in data
- ▶ Access rights to data

Property Rights

- ▶ Rights over the data (nearest thing to “ownership”)
 - ▶ database right
 - ▶ database copyright
- ▶ Rights of access to data (eg API's)
- ▶ **NOT** other forms of intellectual property rights
 - ▶ eg copyright in images, user contributed text
 - ▶ too broad a topic
- ▶ European Union only
 - ▶ EU more protective than US in general

Database Directive (96/9/EC)

a collection of independent works, data or other materials which (a) are arranged in a systematic or methodical way, and (b) are individually accessible by electronic or other means

- ▶ Lots of things are databases (libraries, poetry books,...)
- ▶ Distinguish the database from its contents

Database Right

- ▶ **Substantial investment**
- ▶ in:
 - ▶ obtaining
 - ▶ verifying
 - ▶ presenting
- ▶ Rewards investment
- ▶ “Substantial” can be qualitative as well as quantitative

Football

The Football leagues have much data.

- ▶ Fixtures lists

Arsenal v Man City 12:45 (13 September 2014)

- ▶ Football data

- ▶ when goals were scored
- ▶ who was “man of the match”
- ▶ ...

Football II

- ▶ Fixtures lists
 - ▶ No database right (*Fixtures Marketing* C-444/02)
 - ▶ No independent investment in obtaining, verifying or presenting
 - ▶ Big shock. What does it mean?
- ▶ Football data
 - ▶ Database right (*Football Dataco v Stan James* - Court of Appeal)
 - ▶ Not tested in Europe

Infringement

Infringement tells you the extent of a right. For database right:

- ▶ Extraction (of a substantial part)
 - ▶ i.e. **copying**
 - ▶ occurs if a substantial part of the data appears elsewhere
 - ▶ can be done by hand, need not end up with same schema
- ▶ Re-utilization (of a substantial part)
 - ▶ i.e. making available to the public
 - ▶ generally: publishing on the web
 - ▶ meta search engine (Innoweb v Wegener C-202/12)
- ▶ Repeated and systematic insubstantial extraction and/or re-utilisation
 - ▶ conflicts with normal exploitation of database
 - ▶ or unreasonably prejudice's maker's legitimate interests

Database Copyright

- ▶ An instance of copyright
- ▶ Author's "own intellectual creation" in:
 - ▶ selection
 - ▶ arrangementof the **contents** of the database
- ▶ Not football fixtures lists (Football Dataco v Yahoo! C-604/10)

Example: German Poems

- ▶ 100 most important poems in German literature between 1730 and 1900
 - ▶ significant data mining effort given to
 - ▶ German Professor for final decision
- ▶ Used as inspiration for another list
 - ▶ infringement of database right by extraction (Directmedia Publishing v Albert-Ludwigs-Universitt Freiburg C-304/07)
 - ▶ infringement of database copyright (Federal Court of Justice)

Access Rights

- ▶ Scraping
- ▶ Using robots

Unlawful Access

- ▶ S.1 Computer Misuse Act 1990
 - ▶ intent to secure access
 - ▶ access is unauthorised
 - ▶ defendant knows it is unauthorised
- ▶ Applies to data in or defendant accessing from the UK
- ▶ Scraping not (yet) been prosecuted

Publishing

Dead or alive?

- ▶ 100 greatest poems in German... [Dead]
- ▶ OpenCorporates [Alive]

Options

- ▶ Dead (time insensitive) - licensing
- ▶ Alive (time sensitive) - access control

Licensing

- ▶ Commercial terms
- ▶ Open
 - ▶ attribution
 - ▶ viral
 - ▶ non-commercial

Open Licenses

- ▶ Open Government Licence (not just the public sector)
- ▶ Creative Commons (all works)
 - ▶ CC0
 - ▶ CC-BY (attribution)
 - ▶ CC-BY-SA (viral)
- ▶ Open Data Commons (just databases)
 - ▶ PDDL
 - ▶ ODC-BY
 - ▶ ODbL

Viral Open Licences

- ▶ Creative Commons 4.0 Attribution Share-Alike (CC-BY-SA)
 - ▶ one international licence
 - ▶ now covers databases as well as their contents
- ▶ Open Data Commons Open Database Licence (ODbL)
 - ▶ purely for the database (not contents)
 - ▶ also a contract (belt and braces)
 - ▶ prevents derivative works

Example: OpenCorporates

- ▶ API access to corporate data
- ▶ Free API
 - ▶ Data - ODbL
 - ▶ Must demonstrate contribution to the community
- ▶ Commercial API
 - ▶ Data - without share-alike restriction
 - ▶ Fee for use