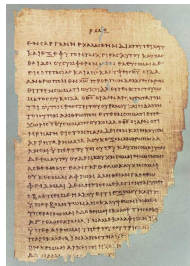


# Understanding law and licensing collecting, using and publishing data

**Francis Davey**

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



# Overview

- ▶ Consumer
  - ▶ Collecting
  - ▶ Using
- ▶ Producer
  - ▶ Publishing

# Scope

Not today: Copyright, for example in:

- ▶ images
- ▶ sounds
- ▶ text (eg user comments)

## 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

# Database Right

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

# Football

- ▶ 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

For all IP rights - infringement tells you the extent of the 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
  - ▶ arrangement
- ▶ Not football fixtures lists (Football Dataco v Yahoo!)

## 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 (German court)

# Access Rights

- ▶ scraping
- ▶ using robots

# 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

# 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)