Derestricting Datasets
How to License Research Data

Alex Ball
DCC/UKOLN, University of Bath

25 February 2011

This work is licensed under Creative Commons BY-NC-SA 2.5
Scotland: http://creativecommons.org/licenses/by-nc-sa/2.5/scotland/

Funded by JISC
Outline

Why license research data?

Licensing options
  Creative Commons
  Open Data Commons
  Open Government Licence
  GILF/AusGOAL Licences
  Design Science Licence
  Public Domain

Mechanisms for Licensing Data

Further Information

Group Exercise
Why license research data?

The open data argument

▶ To allow the data to be used in new ways: comparative studies, data mining, interdisciplinary studies, ‘citizen science’…
▶ To permit greater scrutiny of research
▶ To raise standards of documentation
▶ To protect researchers from challenges
▶ To accelerate community-wide learning from experience
▶ To increase efficiency
▶ To increase impact
Why license research data?

The open data argument

▶ To allow the data to be used in new ways: comparative studies, data mining, interdisciplinary studies, ‘citizen science’…
▶ To permit greater scrutiny of research
▶ To raise standards of documentation
▶ To protect researchers from challenges
▶ To accelerate community-wide learning from experience
▶ To increase efficiency
▶ To increase impact

The pragmatic argument

▶ To provide clarity
Types of licenses

- Contracts

IPR and Licensing: tips, traps and techniques 25 February 2011
Types of licenses

- Contracts
- Pure licences
Types of licenses

- Contracts
- Pure licences
- Waivers
Licensing questions

1. Do you need to make a choice?
   - Institutional policy
   - Data archive policy
2. Would a standard licence suffice?
3. Do you need to write your own licence?
4. Do you need more than one licence?
Multiple licensing

Development Community

Core Product

Development Partners

Copyleft

Licence

Licence fees

Reseller

Licence fees

Copyleft

Licence

Copyleft Users

Customers

Migration

IPR and Licensing: tips, traps and techniques

25 February 2011
Creative Commons

- **BY** Attribution
- **NC** Non-Commercial
- **SA** Share Alike
- **ND** No Derivatives
Creative Commons

- **BY** Attribution
- **NC** Non-Commercial
- **SA** Share Alike
- **ND** No Derivatives

Severely restricts use
Creative Commons

BY Attribution
NC Non-Commercial
SA Share Alike
    Reduces interoperability
ND No Derivatives
    Severely restricts use
Creative Commons

BY  Attribution

NC  Non-Commercial
What counts as commercial?

SA  Share Alike
Reduces interoperability

ND  No Derivatives
Severely restricts use
Creative Commons

- **BY** Attribution
  - Attribution stacking

- **NC** Non-Commercial
  - What counts as commercial?

- **SA** Share Alike
  - Reduces interoperability

- **ND** No Derivatives
  - Severely restricts use
Open Data Commons

- Attribution Licence (ODC-BY)
  - Publicly available

- Open Database Licence (ODC-ODbL)
  - Explicitly distinguishes database structure from contents, and deriving visualisations from deriving new databases.
Open Government Licence

- Attribution 📊
- Terminates on illegal/misleading usage of data
- Can only be used for public sector data
- Cannot be used for logos, insignia, personal data, otherwise encumbered data
- Does not distinguish database structure from contents
GILF/AusGOAL Licences

- Six Australian CC Licences
- Restrictive Licence
  - Contract template: standard legal code modified using schedules
  - Expiry date
  - Geographical restriction
  - Different copying/distribution terms for confidential and ordinary data
  - Licence fees
  - Other restrictions and permissions
Design Science Licence

- Attribution and Share Alike 
- Distinguishes source data from visualisations
- Does not distinguish database structure from contents
- Redistribution requirements
Public Domain

- Creative Commons Zero (CC0)
- Open Data Commons Public Domain Dedication and Licence (ODC-PDDDL)
- Open Data Commons Database Contents Licence (ODC-DbCL)
- Community norms?
[This database is/These data are/(name of dataset) is] made available under the Public Domain Dedication and License v1.0 whose full text can be found at:
http://www.opendatacommons.org/licenses/pddl/1.0/
Attaching the licence to the data

[This database is/These data are/(name of dataset) is] made available under the Public Domain Dedication and License v1.0 whose full text can be found at:

http://www.opendatacommons.org/licenses/pddl/1.0/

http://www.example.com/data/set/1

http://purl.org/dc/terms/license

http://www.opendatacommons.org/licenses/pddl/1.0/
Thank you for your attention

DCC Website: http://www.dcc.ac.uk/
Alex Ball: http://www.ukoln.ac.uk/ukoln/staff/a.ball/
Group Exercise

How would you license the following datasets?

1. Power consumption at your house
2. Temperature and humidity readings from a well-sited Stevenson Screen
4. Anonymised genomic/proteomic data
5. Amounts of carbon and energy embedded in building materials
6. Recordings, transcriptions and annotations of simulated design meetings
Group Exercise

How would you license the following datasets?

1. Power consumption at your house
Group Exercise

How would you license the following datasets?

1. Power consumption at your house
2. Temperature and humidity readings from a well-sited Stevenson Screen
Group Exercise

How would you license the following datasets?

1. Power consumption at your house
2. Temperature and humidity readings from a well-sited Stevenson Screen
How would you license the following datasets?

1. Power consumption at your house
2. Temperature and humidity readings from a well-sited Stevenson Screen
4. Anonymised genomic/proteomic data
Group Exercise

How would you license the following datasets?

1. Power consumption at your house
2. Temperature and humidity readings from a well-sited Stevenson Screen
4. Anonymised genomic/proteomic data
5. Amounts of carbon and energy embedded in building materials
Group Exercise

How would you license the following datasets?

1. Power consumption at your house
2. Temperature and humidity readings from a well-sited Stevenson Screen
4. Anonymised genomic/proteomic data
5. Amounts of carbon and energy embedded in building materials
6. Recordings, transcriptions and annotations of simulated design meetings