



*Citation for published version:*

Ball, A 2011, 'Derestricting Datasets: How to License Research Data', IPR and Licensing: tips, traps and techniques, Centre Point, London, 25/02/11.

*Publication date:*  
2011

[Link to publication](#)

**University of Bath**

**Alternative formats**

If you require this document in an alternative format, please contact:  
[openaccess@bath.ac.uk](mailto:openaccess@bath.ac.uk)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Derestricting Datasets

How to License Research Data

Alex Ball

25 February 2011

## Abstract

This talk is intended to help workshop participants decide how to apply a licence to their research data, and which licence would be most suitable. It covers why licensing data is important, the impact licences have on future research, and the potential pitfalls to avoid.

## Contents

<b>1</b>	<b>Why license research data?</b>	<b>2</b>
<b>2</b>	<b>Licensing options</b>	<b>2</b>
2.1	Creative Commons . . . . .	3
2.2	Open Data Commons . . . . .	3
2.3	Open Government Licence . . . . .	3
2.4	GILF/AusGOAL Licences . . . . .	4
2.5	Design Science Licence . . . . .	4
2.6	Public Domain . . . . .	4
<b>3</b>	<b>Mechanisms for Licensing Data</b>	<b>4</b>
<b>4</b>	<b>Further Information</b>	<b>5</b>
<b>5</b>	<b>Group Exercise</b>	<b>6</b>

# I Why license research data?

The open data argument

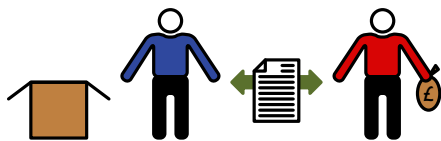
- To allow the data to be used in new ways: comparative studies, data mining, interdisciplinary studies, 'citizen science'...
- To permit greater scrutiny of research
- To raise standards of documentation
- To protect researchers from challenges
- To accelerate community-wide learning from experience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

The pragmatic argument

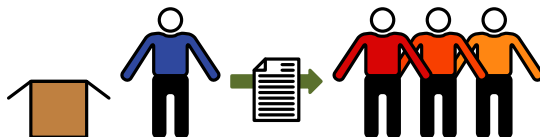
- To provide clarity

## 2 Licensing options

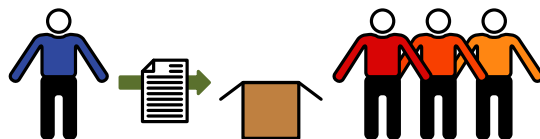
- Contracts



- Pure licences



- Waivers



1. Do you need to make a choice?

- Institutional policy
- Data archive policy

2. Would a standard licence suffice? (See sections 2.1 to 2.6)

3. Do you need to write your own licence?

4. Do you need more than one licence? (See Figure 1)

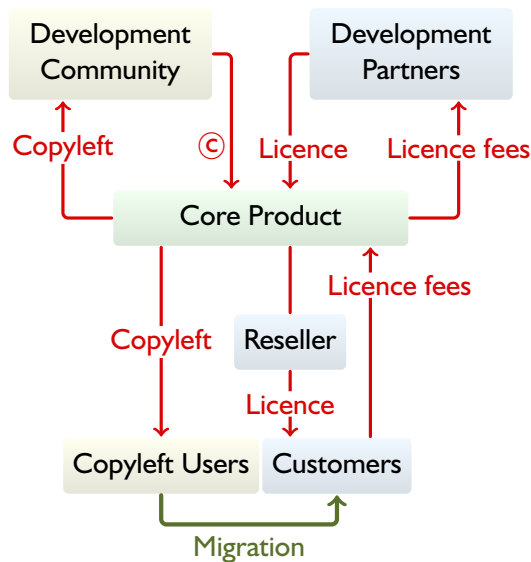


Figure 1: Licence streams of a core product in a simplified dual licensing model. (Välämäki, M. [2003]. Dual licensing in open source software industry. *Systemes d'Information et Management*, 8 [1], 63–75. Retrieved 3 September 2010, from <http://opensource.mit.edu/papers/va1imaki.pdf>)

## 2.1 Creative Commons

- ⓘ BY Attribution  
*Attribution stacking*
- Ⓜ NC Non-Commercial  
*What counts as commercial?*
- Ⓢ SA Share Alike  
*Reduces interoperability*
- ⊖ ND No Derivatives  
*Severely restricts use*



## 2.2 Open Data Commons

- Attribution Licence (ODC-BY)



- Open Database Licence (ODC-ODbL)



- Explicitly distinguishes database structure from contents, and deriving visualisations from deriving new databases.



## 2.3 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



## 2.4 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



AusGOAL = Australian Governments Open Access and Licensing Framework

## 2.5 Design Science Licence

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

## 2.6 Public Domain

- Creative Commons Zero (CC0)
- Open Data Commons Public Domain Dedication and Licence (ODC-PDDL)
- Open Data Commons Database Contents Licence (ODC-DbCL)
- Community norms?

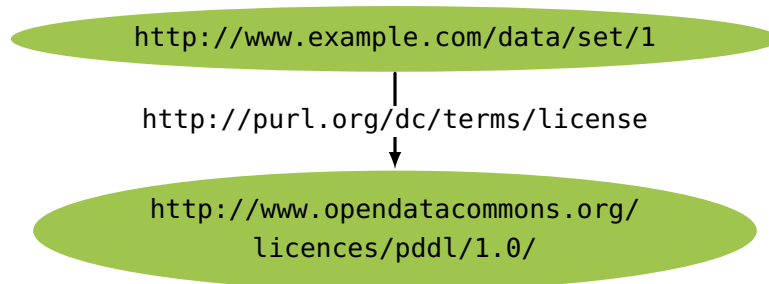


# 3 Mechanisms for Licensing Data

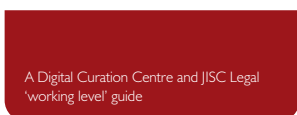
A human-readable statement such as the following should be placed anywhere a user of the data might reasonably be expected to see it, e.g. on the download page, in a readme file, on the document properties.

[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/>

If the data might be harvested automatically, the licence should also be made clear using RDF, e.g.



## 4 Further Information



### How to License Research Data

Alex Ball (DCC)

<http://www.dcc.ac.uk/resources/how-guides/license-research-data>



<http://www.web2rights.com/OERIPRSupport/>

- Starter Pack
- Diagnostic Tools



Digital Curation Centre, 2011.  
Licensed under Creative Commons BY-NC-SA 2.5 Scotland:  
<http://creativecommons.org/licenses/by-nc-sa/2.5/scotland/>



Alex Ball. DCC/UKOLN, University of Bath. <http://www.ukoln.ac.uk/ukoln/staff/a.ball/>



Except where otherwise stated, this work is licensed under Creative Commons BY-NC-SA 2.5 Scotland: <http://creativecommons.org/licenses/by-nc-sa/2.5/scotland/>



The DCC is funded by JISC.

For more information, please visit <http://www.dcc.ac.uk/>

## 5 Group Exercise

How would you license the following datasets?

1. Power consumption at your house. Source: *Andy Stanford-Clark's twittering house* (<http://stanford-clark.com/>); Twitter licence is sort of CC BY, without the BY.
2. Temperature and humidity readings from a well-sited Stevenson Screen
3. Wealth and assets survey data (2006–2008). Source: *ESDS* (<http://www.esds.ac.uk/findingData/snDescription.asp?sn=6415>); UKDA Special Licence.
4. Anonymised genomic/proteomic data. Source: *Sage Commons*; Public Domain.
5. Amounts of carbon and energy embedded in building materials. Source: *University of Bath* (<http://www.bath.ac.uk/mech-eng/sert/embodied/>); no visible licence.
6. Recordings, transcriptions and annotations of simulated design meetings. Source: *AMI Project, University of Edinburgh* (<http://corpus.amiproject.org/>); dual licence – CC BY-NC-SA equiv and Bespoke.