

cessda eric

Consortium of European Social Science Data Archives
European Research Infrastructure Consortium

Copyright, Access Controls and Licensing Data

Dr Scott Summers

UK Data Service

University of Essex

Pravni in etični vidiki ravnanja z
raziskovalnimi podatki - 11th April 2018

Overview

Licensing and access controls can help share sensitive and confidential data in a safeguarded way

- » Areas to be covered:
 - » Copyright
 - » Licensing models
 - » Licensing considerations
 - » Access controls models
 - » Access controls in practice
 - » Access and licence control strategies

Copyright

cessda eric

Copyright

- » Copyright is internationally recognised form of intellectual property right, which arises automatically as a result of original work such as research
- » Copyrighted output from research could include spreadsheets (and other forms of originally selected and organised data), publications, reports and computer programs
- » Copyright will not cover the underlying facts, ideas or concepts, but only the particular way in which they have been expressed
- » The right will lie with the author of the work, or with their relevant institution—different universities will have different policies on intellectual property
- » A copyrighted work cannot usually be published, reproduced, adapted or translated without the owner's permission



Copyright - Key Considerations

- » Questions to ask:
 - » Who the copyright holder of the datasets is?
 - » Are you allowed to use them and in what way?
 - » Are you allowed to archive and publish them in a data repository?
- » Key considerations
 - » Joint ownership
 - » Datasets created by multiple researchers
 - » Derived datasets
 - » Database rights
 - » Provision in a contract
 - » Repository copyright rules



Copyright Scenarios

» **Case Study 1 – Copyright of Archived Data**

- » A researcher uses International Social Survey Programme (ISSP, n.d.) data obtained from ZACAT/GESIS - Leibniz Institute for the Social Sciences in Germany. These data are freely available to registered users. The researcher incorporates some of the ISSP data within a database containing his own research data. Can this database be deposited with another archive?

» **Case Study 2 – Copyright of Data in the Public Domain**

- » A researcher studies how health issues around obesity are reported in the media in the last 10 years. Freely available newspaper websites and library sources are used to obtain articles on this topic. Articles or excerpts are copied into a database and coded according to various criteria for content analysis. (i) Can the researcher use such public data without breaching copyright? (ii) Can the database be archived and shared with other researchers?

Copyright Scenarios

» **Case Study 3 – Copyright of Survey Questions**

- » A researcher wishes to reuse a set of questions from an existing survey questionnaire, to compare results between the newly proposed survey and the original

» **Case Study 4 – Copyright of Interviews with Stay-at-Home Parents**

- » A researcher interviews various stay-at-home parents about their careers and produces audio recordings and near verbatim transcripts herself. The researcher analyses this material and offers it to a data archive. The researcher did not get signed copyright transfers for the interviewees' words. What are the rights issues surrounding this offer of data

Licensing

Publishing Data – Licensing

- » If you publish data with a data repository of your choice, a licence agreement will be applied to your data
- » A licence agreement is a legal arrangement between the creator/depositor of the dataset and the data repository, signifying what a user is allowed to do with the data
- » To make reuse as likely as possible one should use a licence that:
 - » Makes data available to the widest audience possible; and
 - » Makes the widest range of uses possible (as open as possible, as closed as necessary)

Creative Commons Licenses

- » Creative Commons forge a balance inside the traditional “all rights reserved” setting that copyright law creates
- » There are 6 licenses
- » Specifying different requirements
- » All require attribution

creative commons LICENSES

	Copy & Publish	Attribution Required	Commercial Use	Modify & Adapt	Change License
Public Domain	✓	✗	✓	✓	✓
BY Attribution	✓	✓	✓	✓	✓
BY-SA Attribution ShareAlike	✓	✓	✓	✓	✗
BY-ND Attribution NoDerivs	✓	✓	✓	✗	✓
BY-NC Attribution NonCommercial	✓	✓	✗	✓	✓
BY-NC-SA Attrib NonComm ShareAlike	✓	✓	✗	✓	✗
BY-NC-ND Attrib NonComm NoDerivs	✓	✓	✗	✗	✓

Creative Commons Licenses – Notes

» CC0

- » CC0 is a completely open CC licence
- » The copyright owner waives all its rights, including the database right and the right to be identified as the creator

» Notes

- » A CC licence cannot be revoked once it has been issued
- » The licence choice may be limited or determined by the data repository of your choice

In Practice: CC Tool to Help Choose a Licence

Creative Commons - Share your work

Share your work

Use Creative Commons tools to help share your work. Our free, easy-to-use copyright licenses provide a simple, standardized way to give you permission to share and use your creative work— on conditions of your choice. You can adopt one of our licenses by sharing on a platform, or choosing a license below.

Choose a license

This chooser helps you determine which Creative Commons License is right for you in a few easy steps. If you are new to Creative Commons, you may also want to read [Licensing Considerations](#) before you get started.



Choose Features → **Optional Info** → **Get License**

[Get Started](#)

License Features

Your choices on this panel will update the other panels on this page.

Allow adaptations of your work to be shared?


Yes No Yes, as long as others share alike

Allow commercial uses of your work?

Yes No

Selected License

Attribution 4.0 International



This is a Free Culture License!

Help others attribute you!

This part is optional, but filling it out will add machine-readable metadata to the suggested HTML!

Licensing Considerations

- » Rights and ownership, establish who owns what
 - » Questions to ask:
 - » *Were secondary sources used? Are the necessary permissions in place to republish?*
 - » *Were any of the data purchased? What agreement was in place for the future archiving of the data?*
 - » *Any additional copyright considerations?*
 - » *Can you sign a licence on behalf of rights owners?*
- » Sharing more disclosive data may require a data sharing agreement and access procedures
- » Depositor licence agreement
 - » Responsibilities and liabilities
 - » Copyright

Access Controls

Publishing Data – Access Controls

- » Publishing data in a data repository does not automatically make them openly accessible
- » Personal data can still be protected by limiting access to the data
- » Access controls can permit control down to an individual file level, meaning that mixed levels of access control can be applied to a data collection

Access Conditions

» Most data repositories operate a three-tiered approach to data access:

1. Open access

2. Access for registered users (safeguarded)

3. Restricted access

* Embargo

In Practice: Managing Access to Data at the UK Data Archive

Open

- available for download / online access under open license without any registration

Safeguarded

- available for download / online access to logged-in users who have registered and agreed to an End User License (*e.g. not identify any potentially identifiable individuals*)
- special agreements (depositor permission; approved researcher)
- embargo for fixed time period

Controlled

- available for remote or safe room access to authorised and authenticated users whose research proposal has been vetted and who have received training

In Practice: Data with Access Conditions

- » Health and Social Consequences of the Foot and Mouth Disease Epidemic in North Cumbria, 2001-2003 (study 5407 in UK Data Archive collection) by M. Mort, Lancaster University, Institute for Health Research
 - » Interviews (audio and transcript) and written diaries with 54 people
 - » 40 interview and diary transcripts are archived and available for re-use by registered users (**Safeguarded**)
 - » 3 interviews and 5 diaries were embargoed until 2015 (**Safeguarded – Embargoed**)
 - » Audio files archived and only available by permission from researchers (**Safeguarded – Special Agreement**)
- » [Discover.ukdataservice.ac.uk/catalogue/?sn=5407](https://discover.ukdataservice.ac.uk/catalogue/?sn=5407)
- » [Doc.ukdataservice.ac.uk/doc/5407/mrdoc/pdf/q5407userguide.pdf](https://doc.ukdataservice.ac.uk/doc/5407/mrdoc/pdf/q5407userguide.pdf)

In Practice: Access & Licensing ReShare

- » **Use terms and conditions for open access data**
- » Data files deposited as Open Data are licensed under the Depositors choice of one of two Creative Commons Attribution 4.0 licenses:
- » CC-BY-NC-SA Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International
- » CC-BY-SA Creative Commons Attribution-ShareAlike 4.0 International (this licence allows for commercial use)
- » ReShare emphasises that Creative Commons licenses can only be agreed to by the copyright holder(s) and should not be used if there are third party rights holders

The screenshot shows the 'Add a new file or zip bundle' form in the ReShare system. It includes instructions on uploading data files and documentation, and a list of fields for configuring the file's access and licensing.

Add a new file or zip bundle

Upload data files and documentation for your data collection. Click the Browse button below to select a file, and the Upload button to upload it to the repository. Large files (>2GB) may take a long time to upload, depending on your browser speed. When uploading many files (>10), please upload them in logically grouped zip bundles instead, e.g. grouping by content (data, documentation), by file format (text, images, databases) or by folder structure.

Make sure data files contain no dedictive personal information.

Include documentation on data collection methods and context to inform future reuse of the data.

Make sure file names contain no spaces.

File: From URL:

Choose File: no file selected

Interview_Data_001.docx
200k

Hide options

Accessible to: Selecting an access level will determine who is able to download this file or bundle:

- open data are accessible to any user without registration
- safeguarded data are accessible only to users registered with the UK Data Service

Consult our detailed [data use terms and conditions for open data and safeguarded data](#). If you think that more access controls to the data are needed, please provide information to the ReShare administrator in the "Notes to administrator" text box (previous page).

Anyone (open access data)

File or bundle content: Data

File or bundle description:

Embargo date:

License: Specify an explicit license for this file or bundle. This repository allows Creative Commons licenses for open data (accessible to anyone) and the UK Data Service End User License for safeguarded data accessible to registered users of the UK Data Service only.

Creative Commons Attribution 4.0 International

Save

< Previous Save for later Cancel Next >

Access Control Strategy

- » When choosing an access category, consider the following:
- » Does the data contain identifiable information?
- » Can the information in this data collection be linked with anything in another data collection which might lead to participant's identities being disclosed?
- » What did participants consent to?
- » If 'restricted access' is to be chosen who will manage the access to this request?

Strategy for Enabling Safe Access

» Fives safes enables safe access to data that meet the needs of data protection. Yet fulfils the demands for open science and transparency

1. Safe data - treat data to protect confidentiality
2. Safe people - educate researchers to use data safely
3. Safe projects - research projects for 'public good'
4. Safe settings - Secure Lab system for sensitive data
5. Safe outputs - Secure Lab projects outputs screened

» [5 Safes Animation](#)



Open Metadata for Sensitive Data

- » Even if data cannot be published in open access, it is always possible to publish the metadata which belongs to the dataset
- » Openly publishing metadata is the only way to make such datasets discoverable
- » Metadata is always freely accessible meaning that:
 - » No registration is needed for searching in the metadata;
 - » No registration is needed for harvesting the metadata (e.g. by search engines)
- » Metadata of sensitive datasets should never contain confidential or identifying elements or characteristics, like names
- » When someone finds a dataset under restricted access, they can submit an access request to the dataset holder. If this is granted, the dataset will be available to download by the user

Concluding Remarks

- » Sensitive and confidential data can be safeguarded by regulating or restricting access to – and use of – the data
- » Access controls should always be proportionate to the kind of data and level of confidentiality involved
- » When regulating access and licensing data, consider who would be able to access the data, what they are able to do with it, whether any specific use restrictions are required, and for how long the data are to be available for

Questions

- » Dr Scott Summers
- » Collections Development and Producer Relations team
- » UK Data Service
- » University of Essex
- » ukdataservice.ac.uk/help/get-in-touch