

Introduction to copyright: Copyright issues in secondary data use

Dr Hina Zahid

Senior Research Data Officer

25 May 2023



Overview

- What is meant by Intellectual Property Rights (IPR)?
- What is secondary data?
- What rights might there be in research data?
- Special consideration.
 - Licensing.
 - Who owns the right?
 - Social media data.
 - Copyright in the international context.
- Resources.

What is meant by Intellectual Property (IP) rights?

- **IP rights:** Rights granted to creators and owners of works that are the result of human intellectual creativity.
- They include:
 - **Trade marks** which distinguish goods and services.
 - **Patents** for new inventions.
 - **Registered designs** for the design of products.
 - **Copyright** for creative works such as books, paintings and music.

What is secondary data?



A quick question

Have you used or plan to use secondary data?

What rights might there be in research data?

Copyright



Database rights

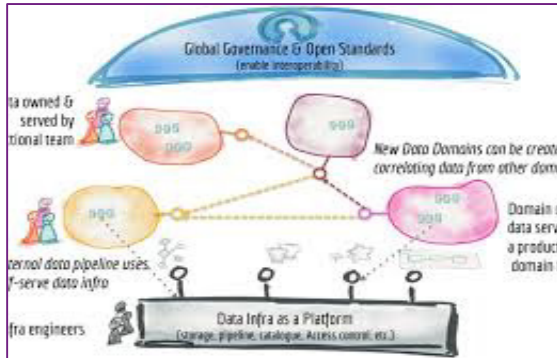
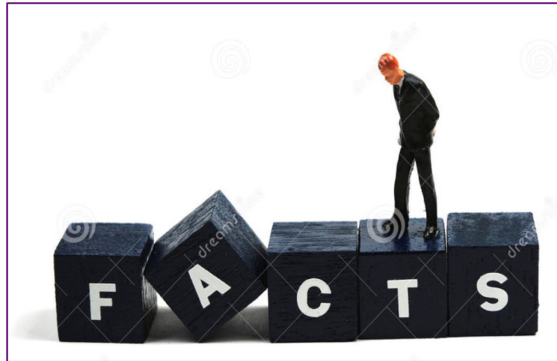


Further information on [rights in research data](#) and [different scenarios](#).

Copyright

- Copyright is an intellectual property right assigned automatically to the creator.
- Applies to (UK Copyright, Designs and Patents Act 1988):
 - original literary, dramatic, musical or artistic works.
 - sound recordings, films, broadcasts or cable programmes.
 - the typographical arrangement of publications.

When my research data be protected by copyright? contd..



Month	T _a (°C)	GHI (W/m ²)	DHI (W/m ²)
January	6.8	116	67
February	11.5	182	106
March	13.1	212	127
April	19.4	285	182
May	19.9	302	214
June	25.5	349	269
July	25.1	344	265
August	24.9	299	227
September	22.3	262	194
October	20.2	196	135
November	14.2	150	99
December	7.5	109	68



What actions require permission?

- Permission is not required:
 - Incorporation of the factual data in one's own words in a structure of one's own.
 - Making a copy (including a digital copy) and utilizing that copy for one's own research, as long as it is not made available to others.
 - Citing from the research data.
- Permission is required:
 - Inclusion of the research data in a publication.
 - Sharing the research data with other people.
 - Including the whole of the research data in a database of one's own that is also shared with other people.

Copyright and personal data

Where personal data is concerned:

- Permission from the person who has created the work is required (the copyright holder).
- Permission from all the people whose personal data is in the work is required. This includes opinions about people.

Special considerations

- Licensing Framework.
- Who owns the right in research?
- Social media data.
- Copyright in the international context.

When my research data be protected by copyright?

When data is created, shared & reused



Data Licencing

Data collections can be made available under:

- Open licenses.
- Bespoke licenses.

Example of open licenses in the Creative Commons Framework:

License	Use and Share	Attribute the work (copyright)	Use commercially	Adapt/create derivatives	Change license
CC BY	✓	✓	✓	✓	✓
CC BY-SA	✓	✓	✓	✓	X
CC BY-NC	✓	✓	X	✓	✓

Bespoke Licencing

Most of the data made available via responsible repository is made available under Bespoke Licenses due to the residual risk of disclosure in data.

The conditions of use require researchers to **act responsibly and ethically** with the data.

[UK Data Service End User Licence Agreement \(EUL\)](#).

When using secondary data always ensure you are familiar with the terms and conditions under which the data is made available.

UKDS Licencing Framework

Open

- No real disclosure risk. Under open licences; almost no restrictions on reuse

Safeguarded

- Zero to low real disclosure risk. Requires authentication and authorisation e.g. registered user and End User Licence Agreement

Controlled

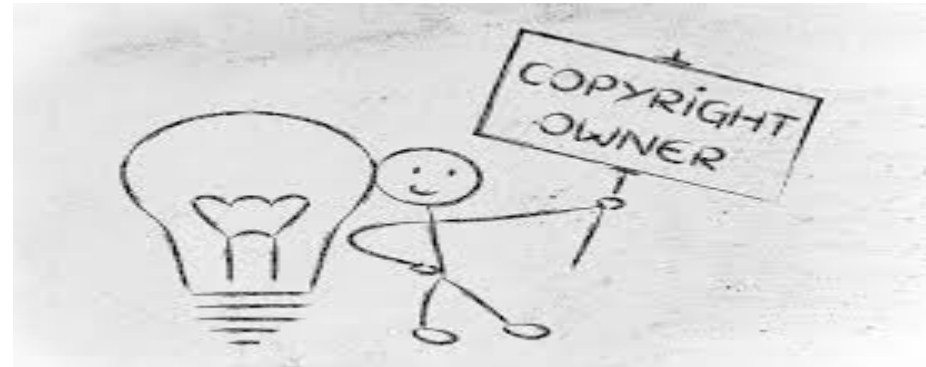
- Real disclosure risk. Requires project approval, user vetting and training; access via a safe setting; output checking

Who owns the right in research?

Special considerations!!!

Who owns the right?

- University or an employee?
- University or a student?
- Research funder or researcher?
- Collaborators in a research project?



- **Best practise is to find out the ownership ASAP.**

How to find out who owns the IP rights?

- The applicable national IP law.
- IP policies of the university/PRI.
- The individual contractual agreements among the university/PRI, creators and sponsors.
- Legal advise.
- Failure to clarify rights can lead to unexpected limitations to:
 - your research.
 - its dissemination.
 - future related research projects.
 - associated profit or credit.

What is social media?

Internet-based or mobile applications which allow users to form online networks or communities



Social media data

- Usually obtained through the application programming interfaces (APIs) of the social media platforms (directly or via resellers).
- What does social media data includes?
 - Individual posts or tweets.
 - Comments.
 - Photos or videos.
 - Networking connections or social interactions.
 - Data on ratings.
 - Interests.
 - Likes or dislikes.

Copyright and social media data

- Content is protected by copyright in the same way as books and journals.
- Users have copyright for their own content (e. g. Tweets or Facebook posts).
- Agreement between the user and the social media platform allows a site to freely use the work for a variety of purposes.
- Researchers need to abide by the terms and conditions of the platforms or API developers.

Challenges when sharing social media data: Twitter

- Sharing of data is prohibited under Twitter's API terms of service.
- Tweet IDs associated with each tweet, and account IDs associated with each user can be shared.
- Data recreation from an archived twitter data can be a real challenge.
- Content modification is not allowed, so the content cannot be anonymised.

Useful checklist (Source: UCL)

- Has Twitter been informed of the proposed work through the Developer API signup mechanism and has it agreed?
- Is any access to Twitter proposed that does not use the API for retrieval (e.g. prohibited methods like scraping)?
- Has the research team read and understood all of the terms that apply to their work?
- Is data to be retained locally? If so, how will this be synchronised sufficiently frequently to reflect the live Twitter data?
- Is data to be shared? If so, does the research protocol for doing so comply with the terms of access?
- How is the informed consent of Tweeters to be gained (or what argument is being used to justify waiving this requirement)?
- Does the informed consent of Subjects need to be considered and if so, how has it been?
- Is publication of individual Tweets anticipated? If so, what forms, process, and information will be used to seek informed consent for publication from Tweeters (and if need be, Subjects mentioned in the Tweets)?
- Is the intended analysis likely to derive or infer information in the restricted use cases?
- Are any other restricted activities (off-Twitter matching, profiling etc) proposed? If so, how is this justified to balance the potential breach of contract?
- Has a DPIA been undertaken and data protection registration put in place?
- Does the work qualify as profiling under the GDPR?
- Has the safety of the researchers been appropriately considered given the topic and methods?
- Where a proposal being reviewed appears to be close to the edge of what is permitted, has legal advice been sought on contract compliance?
- Is there a reputational risk to UCL through potential breach of contract?

Copyright in the international context

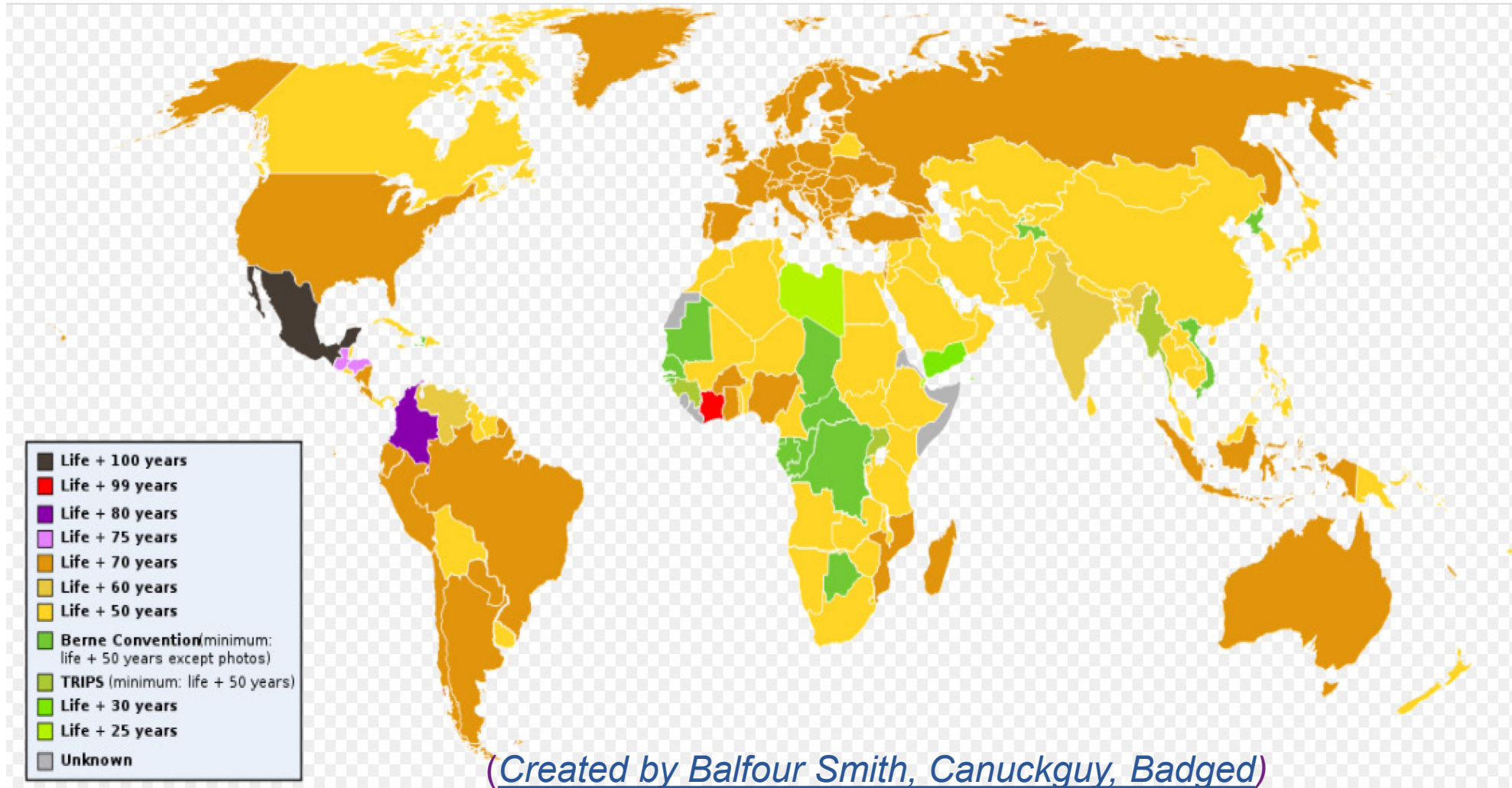
Berne Convention (1886)

- Oldest and most important treaty, signed in **1886**.
- 179 countries covered under Berne Convention.
- Establishes minimum standards of protection.
 - Types of works protected.
 - Duration of protection.
 - Scope of exceptions.
 - Limitations.
 - Principles such as “**national treatment**”.
 - Principles such as “**automatic protection**”.

National variation in copyright law

- Country specific laws but built on International principles.
- Consensus on basic issues, but differences in:
 - Types of works covered (e.g. data sharing).
 - Duration (minimum terms set as life + 50 years).
 - Exceptions (fair dealing).

National variation in copyright law (e.g, differences in copyright duration)



Best practices when using secondary data

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?
-
- Researchers are allowed to copy whole work/data for their own analysis but they may NOT be allowed to share it.
 - Majority of uses of copyright materials continue to require permission.
 - Exceptions to copyright are country specific.

Copyright scenario

A researcher has used secondary data sources for a research project. He plans to share his data. The sources he used were:

- World Bank open data.
- Microsoft Academics.

World Bank open data

Licenses

You are encouraged to use the Datasets to benefit yourself and others in creative ways. You may extract, download, and make copies of the data contained in the Datasets, and you may share that data with third parties according to these terms of use.

Unless specifically labeled otherwise, these Datasets are provided to you under a Creative Commons Attribution 4.0 International License (CC BY 4.0), with the additional terms below. The basic terms may be accessed [here](#). When you download or use the Datasets, you are agreeing to comply with the terms of a CC BY 4.0 license, and also agreeing to the following mandatory and binding addition:

Microsoft Academics

Personal and Non-Commercial Use Limitation

Unless otherwise specified, the Services are for your personal and non-commercial use. You may not modify, copy, distribute, transmit, display, perform, reproduce, publish, license, create derivative works from, transfer, or sell any information, software, products, or services obtained from the Services.

Resources

- [Copyright \(UK Data Service\)](#)
- [Access levels/Licenses \(UK Data Service\)](#)
- [Variable information log template for secondary data users](#)

Variable Information Log									
Variable name	Variable label	Source	Dataset version (if applicable)	URL/DOI	License information	Unit of analysis	Date data downloaded/obtained	Brief description of the data	Data collection method

*Please insert as many rows as needed.

Resources contd...

- [Copyright Fact sheet](#)
- [Working with volunteers: Copyright considerations](#)
- [Copyright and social media](#)
- [Copyright and open licensing- A guide](#)
- [CESSDA Data Management expert Guide](#)
- [Copyright \(UK Data Service\)](#)
- [Access levels/Licenses \(UK Data Service\)](#)
- [Variable information log template for secondary data users](#)

Resources contd..

- [Orphan works diligent search guidance](#)
- [Copyright Scenarios \(UK Data Service\)](#)
- [Copyright Resources \(UK Data Service\)](#)
- [Other rights \(UK Data Service\)](#)
- [Useful checklist for using twitter data](#)(Source: UCL)
- [ONS checklist for social media data](#)
- [Berne Convention](#)
- [Salient features of copyright law in the EU](#)
- CESSDA Data Management Expert Guide: [European Diversity in copyright](#)

Thank you.

Dr. Hina Zahid

hzahid@essex.ac.uk

datasharing@ukdataservice.ac.uk

Twitter: @hinazahid2016

@UKDataService