Management of research data: intellectual property rights and privacy issues

April 6, 2017

Paolo Guarda

Faculty of Law – University of Trento
Image courtesy of jscreationzs at FreeDigitalPhotos.net
Goals of the seminar
RESEARCH DATA
Several definitions

“Research data are factual records, which may take the form of numbers, symbols, text, images or sounds, which are used as primary sources for research, which are commonly accepted in the research community as necessary to validate research findings”

(The Australian Griffith University)
OPEN DATA
What Open Data is?

- Conceptual framework provided by the Open Access movement
  - “Open access (OA) literature is digital, online, free of charge, and free of most copyright and licensing restrictions” P. Suber, 2012

- Open data → digital information available on the Net for free and without (most of) the technological and legal limitations (IPR)
BIG DATA
What is Big Data?

“Datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze”, McKinsey, 2011

• Quantitative aspect → only part of the phenomenon
• Innovative aspect → combination and mining of different data sources (data sylos)
DATA MANAGEMENT PLAN
DMP

• A DMP describes the data management life cycle for the data to be collected, processed and/or generated by a Horizon 2020 project.

• Projects participating in the Horizon 2020 Open Research Data Pilot will be required to develop several versions of a DMP, in which they will specify what data will be kept for the longer term.
  – Other projects are invited to submit a DMP if it is relevant for their planned research.

• Guidelines on FAIR Data Management in Horizon 2020 (v.3 July 2016):
Legal issues involved?

• Intellectual property rights

• Privacy and Data protection

• (Law of contract (Linceses))
Intellectual Property Rights
Is a single data protectable?

May it be subject of Intellectual property rights??

As usual, the more correct answer from a legal point of view is... it depends...!
No protection to data itself

“Facts are not protectable by copyright law [because] are not "original" to any author, but are rather "discovered" by authors”. Samuelson, 1994

• Unless it constitutes an copyrighted work (i.e, a photo)
There’s always a but...
COPYRIGHT
Technology and market

- Before the invention of movable type printing, the original costs the same as the copy (the cost is the payment of the slave, the scribe).

- With movable types, the original (matrix) is very expensive, the copy is cheap (the marginal cost of producing the copies is low).

- The second printer copies the matrix (supporting costs) and then he must support only the marginal costs of producing copies.

- With a new market (the book) and a new business model (selling large-scale copies of the originals with a cover price) comes the need for new rules.
Information as a “public good”

- Information features:
  - Immateriality
  - Inexhaustible
  - incremental and cumulative nature (“on the shoulders of giants”)

- Information → “Public Good”:
  - Not rival
  - Not excludable

- A market of public goods → Market Failure
Remedies to the market failure

- State remedies for market failure:
  - Direct intervention
  - Awards and grants
  - **Monopoly/Property rights** (copyright and patents)
    - To create artificially (ie, by the force of law) the excludibility that is missing to information in the state of nature
    - An exclusive right that allows to apply an higher price with respect to the marginal cost in order to incentive the creation and distribution
    - A poised balance → The exclusive right is limited in time and extent
The Origins of Copyright

- Relationship between the Gutenberg invention of the press and the legal protection of literary works:
  - Press invented in the 15th century: first Bible published in 1455
  - Privilege of the Republic of Venice (1469)
Copyright Law

- Copyright law grants authors the **exclusive rights** in their **intellectual works**.

- The exclusive right embraces
  - Moral rights
  - Economic Rights

- Limits of the exclusive right:
  - **Extension** (originality; expression/idea dichotomy)
  - **Duration**
Copyright Law limits: extension

- **Extension:**
  - Originality
  - Expression/Idea Dichotomy
  - First Sale Doctrine (Once a work is sold or distributed on a specific territory with the consent of the right holder, the latter may not control or prevent the further distribution).
Copyright Law limits: duration

- Statute of Anne – 1710: 14 years (+ 14)
- U.S. Copyright Act 1790: 14 years
- Berne Convention art. 7: author's life + 50 years
- Directive 93/98/EC, 29 October 1993: author’s life + 70 anni
  - Art. 25, l. 22 aprile 1941, n. 633 (Italian) “Copyright Law”
- Sonny Bono Copyright Extension Act of 1998: author’s life + 70 years
SUI GENERIS RIGHTS ON DATABASES
Models of databases legal protection
The legal protection of databases

USA Copyright

• Copyright Act: protection of compilations
  – 17 U.S.C. § 103
• Sweat of the brow doctrine
• Feist v. Rural telephone (1991)

Euro-Italian copyright

• Two-tier system
  – copyright
  – sui generis right
• Directive 96/9/EC
• L. 633/41
The legal protection of databases: EU-ITA Copyright

“databases which, by reason of the selection or arrangement of their contents, constitute the author's own intellectual creation shall be protected as such by copyright”.

Art. 3, Directive 96/9/EC
## Two-tier system of Directive 96/9/EC

<table>
<thead>
<tr>
<th><strong>Copyright</strong></th>
<th><strong>Sui generis right</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Database which “by reason of the selection of arrangement of their contents, constitute the author’s own intellectual creation”</td>
<td>Any database where “there have been qualitatively and/or quantitatively substantial investments in either the obtaining, verification or presentation of the contents”</td>
</tr>
<tr>
<td><strong>Author</strong></td>
<td><strong>Maker of the database</strong></td>
</tr>
<tr>
<td><strong>“expression” of the database</strong></td>
<td>No minimum standard of creativity</td>
</tr>
<tr>
<td>• originality of its systematic organization</td>
<td></td>
</tr>
<tr>
<td>• no mere alphabetical or chronological order</td>
<td></td>
</tr>
<tr>
<td><strong>Duration: 70 years after author’s death</strong></td>
<td><strong>Duration: 15 years (see art. 10.3)</strong></td>
</tr>
</tbody>
</table>
Powers of the maker of the database to prevent *extraction* and/or *re-utilization* of the *whole* or of a *substantial part*, evaluated *qualitatively and/or quantitatively*, of the contents of that database, if the obtaining, verification or presentation of the contents shows that there has been *qualitatively and/or quantitatively* a *substantial investment*

Art. 7, Direttiva 96/9/CE
What does “extraction” mean?

“the permanent or temporary transfer of all or a substantial part of the contents of a database to another medium by any means or in any form” (art. 7 § 2 lett. a)

= L. 633/41, art. 102 bis, para. 1, lett. b)
What does “re-utilization” mean?

“any form of making available to the public all or a substantial part of the contents of a database by the distribution of copies, by renting, by on-line or other forms of transmission” (art. 7 §. 2 lett. b)

= L. 633/41 art. 102 bis, para 1, lett. c)
LICENSES: A BRIEF OUTLINE
Exploitation and contractual tools

- **Assignment**
  - → full and final transfer

- **License**
  - Limited and temporary transfer
    - **Exclusive license** (no person or company other than the named licensee can exploit the relevant intellectual property rights)
    - **Non-exclusive license**: (grants to the licensee the right to use the intellectual property, but means that the licensor remains free to exploit the same intellectual property and to allow any number of other licensees to also exploit the same intellectual property)
Business Models and Licenses

- **Hierarchical Model**
  - Based on traditional copyright
  - Proprietary License (i.e. EULA)

- **Not Hierarchical Model**
  - Peers play hybrid roles (i.e. software)
  - Open Licenses
Not Hierarchical Model: open logic

Hierarchical Model: fixed roles
GNU GPL logics

Licenses by Name

The following licenses have been approved by the OSI via the License Review Process.

- Academic Free License 3.0 (AFL 3.0)
- Affero GNU Public License
- Adaptive Public License
- Apache License, 2.0
- Apple Public Source License
- Artistic license 2.0
- Attribution Assurance Licenses
- New and Simplified BSD licenses
- Boost Software License (BSL1.0)
- Computer Associates Trusted Open Source License 1.1
- Common Development and Distribution License
- Common Public Attribution License 1.0 (CPAL)
- CUA Office Public License Version 1.0
- EU DataGrid Software License
- Eclipse Public License
- Educational Community License, Version 2.0
- Eiffel Forum License V2.0
- Entessa Public License
- European Union Public License (link to every language's version on their site)
- Fair License
- Framework License
http://creativecommons.org/videos/wanna-work-together
LEGAL TOOLS FOR OPEN DATA

licenses

- CC by-sa¹
- ODC ODbL²
- IODL 1.0 (it)³

requiring attribution and share-alike

- CC by⁴
- ODC by⁵
- OGL (uk)⁶
- IODL 2.0 (it)⁷
- LIP (fr)⁸

requiring attribution only

public domain waivers

creating an “artificial” public domain*

- CC0⁹
- ODC PDDL¹⁰

¹) Creative Commons Attribution - Share Alike 3.0 (http://creativecommons.org/licenses/by-sa/3.0/)
²) Open Data Commons Open Database License 1.0 (http://opendatacommons.org/licenses/odbl/1-0/)
³) Italian Open Data License 1.0 (http://www.formez.it/iodl/)
⁴) Creative Commons Attribution 3.0 (http://creativecommons.org/licenses/by/3.0/)
⁵) Open Data Commons Attribution License (http://opendatacommons.org/licenses/by/)
⁷) Italian Open Data License 2.0 (http://www.dati.gov.it/iodl/2.0/)
⁸) Licence «information publique librement réutilisable» v.1 (www.rip.justice.fr/information_publique_librement_reutilisable)
⁹) CC0 1.0 Universal - Public domain dedication (http://creativecommons.org/publicdomain/zero/1.0/)
¹⁰) Open Data Commons Public Domain Dedication and License (http://opendatacommons.org/licenses/pddl/)

*natural (and real) public domain comes when all exclusive rights are expired

1) Creative Commons Attribution - Share Alike 3.0 license

original url for this document: 
www.aliprandi.org/doc/opendata_graph.pdf

a diagram by
Simone Aliprandi
(www.aliprandi.org)
release: 21/05/2012

www.lawtech.jus.unitn.it
DATA PROTECTION REGULATION
European Union Legal Framework

- **Directive 95/46/EC** of the European Parliament and the Council of 24 October 1995 on the Protection of individuals with regard to the processing of personal data and on the free movement of such data


  - Amended in 2009
Main features

- Definition of general principles with regard to the processing
- Acknowledgment of detailed data subject rights
- Specific regulation of the so called “sensitive data”
Italian Data Protection Code

- Legislative Decree 30 June 2003, n. 196 (Data Protection Code)
  - gathers up all the old Italian acts on data protection
  - implements Directive 2002/58/EC
  - establishes new rules in a systematic way
  - divided in 3 main sections
But what is a processing?
Broad legal definition

"'processing of personal data' ('processing') shall mean any operation or set of operations which is performed upon personal data, whether or not by automatic means, such as collection, recording, organization, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction”

(art. 2, lett. a) EU Directive)
Main principles

• Fair and lawful processing
  – Compliance with the law and good faith principle
  – Legitimate grounds for collecting and using the personal data

• Purpose limitation
  – “collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes” (art. 6, par. 1, lett. b)
  – Control of the information flow
  – Problem of secondary uses
Main principles

• Minimal disclosure
  – Privacy by design
  – Privacy by default

• Accurate and, where necessary, kept up to date

• Adequate, relevant and not excessive
Informed consent

• “the data subject's consent' shall mean any freely given specific and informed indication of his wishes by which the data subject signifies his agreement to personal data relating to him being processed”
  (art. 2, lett. h, Directive 95/46/EC)
• Opt –in
• Withdrawable at any time.
• Expression: documented in writing
  – Consent shall be given in writing if the processing concerns sensitive data
Personal data

“shall mean any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity”

Art. 2, lett. a), EU Directive
Sensitive data

“shall mean personal data allowing the disclosure of racial or ethnic origin, religious, philosophical or other beliefs, political opinions, membership of parties, trade unions, associations or organizations of a religious, philosophical, political or trade unionist character, as well as personal data disclosing health and sex life”

(art. 4, co. 1, lett. d, IDPC)
Anonimous data

“data that cannot be associated to any identified or identifiable data subject”
(art. 4, para. 1, lett. n, IDPC).
‘communication’ shall mean disclosing personal data to one or more identified entities other than the data subject, the data controller’s representative in the State’s territory, the data processor and persons in charge of the processing in any form whatsoever, including by making available or interrogating such data (art. 4. para. 1, lett. I), Italian Data Protection Code)
Dissemination

‘dissemination’ shall mean disclosing personal data to unidentified entities, in any form whatsoever, including by making available or interrogating such data”

(art. 4. para 1, lett. m) Italian Data Protection Code)
Who are the main actors?

Data controller
- Decides alone, jointly or in common with another how and why data are processed.

Data Subject
- the person to whom personal data refer

Data Processor(s)
- Processes data under data controller’s instruction.
New framework: GDPR

- REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)
  - Published: 4 May 2015
  - Applied from: 28 May 2018
Privacy and Scientific Research
Two main phases in the dp cycle

• **Collection and use of personal data**
  - Consent + information notice
    – Specific purpose

• **Dissemination of research results containing personal data**
  - Only as aggregate data or in ways that the data subject cannot be identified
Legal framework – EU DPD

• Art. 6, par. 1, lett. b Dir. 95/46: “Member States shall provide that personal data must be: (…) b) collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes. Further processing of data for historical, statistical or scientific purposes shall not be considered as incompatible provided that Member States provide appropriate safeguards;”

• Recital 29: “Whereas the further processing of personal data for historical, statistical or scientific purposes is not generally to be considered incompatible with the purposes for which the data have previously been collected provided that Member States furnish suitable safeguards; whereas these safeguards must in particular rule out the use of the data in support of measures or decisions regarding any particular individual;”

• Recital 40
Legal framework – GDPR

• Art. 89 (Safeguards and derogations relating to processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes), in part. par. 2: “Where personal data are processed for scientific or historical research purposes or statistical purposes, Union or Member State law may provide for derogations from the rights referred to in Articles 15, 16, 18 and 21 subject to the conditions and safeguards referred to in paragraph 1 of this Article in so far as such rights are likely to render impossible or seriously impair the achievement of the specific purposes, and such derogations are necessary for the fulfilment of those purposes”
Legal framework – GDPR

• Recital 33: “It is often not possible to fully identify the purpose of personal data processing for scientific research purposes at the time of data collection. Therefore, data subjects should be allowed to give their consent to certain areas of scientific research when in keeping with recognised ethical standards for scientific research. Data subjects should have the opportunity to give their consent only to certain areas of research or parts of research projects to the extent allowed by the intended purpose”

• …
Legal Framework: Italy

• Italian Data Protection Code:
  – Title VII *Processing for historical, statistical or scientific purposes*
    • In part. Chapter 4 (Processing for statistical or scientific purposes), artt. 104-110
  – Annex 4 – Processing of personal data for statistical and scientific purposes
  – General Authorization of the Data Protection Authority
Main principles

• Primary and secondary uses
• Data can be stored for longer periods
• Specific rules with reference to information notice and consent
• Draft and deposit of a project (Italian context requirement)
• Dissemination (no identification data)
• Data communication among universities and research bodies
• …
Privacy vs. Open Data movement

Specific purpose vs Any purpose
Sum-up

• Several phenomena and key issues involved:
  – Research data, Big data, open data, DMP, etc.

• Several legal issues involved:
  – Copyright and sui generis rights
    • Licenses
  – Data protection and privacy
    • EU regulation
    • Specific rules for scientific research
Thank you

Image courtesy of Stuart Miles at FreeDigitalPhotos.net
paolo.guarda@unitn.it

Webpage:
http://www5.unitn.it/People/it/Web/Persona/PER0003411#INFO
http://www.lawtech.jus.unitn.it/index.php/people/paolo-guarda

This presentation is licensed under a
Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International
<http://creativecommons.org/licenses/by-nc-nd/4.0/>