

Privacy and anonymization in repositories of grey literature

Michal Koscik
Masaryk University

Google vs. Spain

- Spanish citizen, Mr. Gonzalez sued newspaper for posting his personal information online
- The information was truthful and article did not breach any rules at the time of publishing
- The article was old, the information damaged mrg. Gonzalez in the long run
- Google indexed contested decision via google search engine
- The newspaper and google are now forced to remove the content form the website and indexing

Right to be forgotten if the information is

Inaccurate

Inadequate

Irrelevant

Excessive

THE COURT SAID THAT.....

- **finding information published or placed on the internet by third parties,**
- **indexing it automatically,**
- **storing it temporarily**
- **making it available to internet users according to a particular order of preference must be classified as ‘processing of personal data’**

... is data processing

Consequences for grey literature ?

Repository

Search engine

Is there any personal data in grey literature?

- Mainly accidental
 - Annual reports, contact information,
 - Websites and blogs
 - Databases
 - Poorly anonymized research data
 - Diploma theses'
 - Photographs
 - Biographies,

**WHAT WOULD YOU DO,
IF SOMEONE ASKED YOU TO REMOVE DOCUMENT
FROM YOUR REPOSITORY ???**

Would you:

Remove the document ?

Anonymize the information ?

Remove keywords form search tool ?

Justify your right to publish the information ?

Basic rule

- European charter of human rights:
 - must be processed fairly for specified purposes
 - on the basis of the consent of the person concerned or some other legitimate basis
 - Everyone has the right of access to data which has been collected concerning him or her,
 - Everyone has the right to have it rectified.

When to:

Remove the document

- Completely excessive for the purpose of the repository

Anonymize the information

- The document is relevant, but the personal information is inadequate, inaccurate or irrelevant

Remove as a keyword from the search site

- Personal information excessive, the document can be still indexed but not found under certain keyword

Justify the use

- See next slide

Justification

the data subject has unambiguously given his consent

- can be taken back

processing is necessary for the performance of a task carried out in the public interest

- Library? University?

processing is necessary for the purposes of the legitimate interests pursued by the controller,

Liability for DP violation

The author of the document

- always

Repository

- Can be protected if it was uploaded by third party

Recommendations

- Establish data removal (privacy) policy
 - Full institution information
 - Contact person
 - Format of request
 - Define the purpose of the database/search tool
 - Indicate public interest
 - Indicate your sources

Thank you for Your attention