

# Speech Data and Privacy Regulation in Europe

Sari Bruun  
Martti Vainio

University of Helsinki



# Contents

- Background
  - Privacy and legislation
  - Data protection and legislation
- Definition: Personal Data
- Guidelines for collecting and using Speech Data
- Exception in Legislation - Scientific Research



# Privacy

- What is privacy?
  - What does it mean to you?
  - Different meanings for different people.  
(eg. teens: something parents and teachers don't see/hear)
  - Even definitions in legislation vary/not unambiguous.



## Privacy Legislation

In Europe the privacy has been recognized as a fundamental right already for decades, e.g.:

- Universal Declaration of Human Rights 1948
- European Convention on Human Rights 1950

Regulated also by national legislation eg.:

- Fundamental laws
- Criminal laws



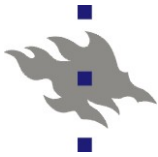
## Personal Data Legislation

- The protection of personal data was, as a separate right, granted to an individual for the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (1981).
- The European Data Protection Directive (Directive 95/46/EC)
- Directive on Privacy and Electronic Communications (Directive 2002/58/EC)



# Personal Data Legislation

- Directives are implemented in EU member states through the national laws (status of implementation: [http://ec.europa.eu/justice\\_home/fsj/privacy/law/implementation\\_en.htm](http://ec.europa.eu/justice_home/fsj/privacy/law/implementation_en.htm))
  - Some variations in implementation.
  - Even more variations in interpretation.
    - Latest example Google Street View –case
      - Germany: Google has to erase the raw footage of faces, house numbers, car license plates and individuals who don't want their information to be used in the service.
      - Greece: Photographing is not allowed at the moment on the streets for Google.
      - Switzerland (non-EU state): Faces and license plates should be made sufficiently unrecognizable, pictures of enclosed areas (e.g. private gardens) should be removed. Court case?



## Personal Data

- Defined in legislation (Data Protection Directive 95/46/EC)
- Personal data:
  - Any information relating to an ***identified or identifiable*** natural person.
  - 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.
  - Wide interpretation.



## Personal Data and Speech Data

- Speech Data typically includes data, which is:
  - ***identified or identifiable***
    - ***directly or indirectly***
      - Recital 26 of the Directive (95/46/EC): whether or not the individual is nevertheless identifiable will depend on "all the means likely reasonably to be used either by the controller or by any other person to identify the said person".
      - Compare: Face to face –contacts/others.
- Legislation is technologically neutral:
  - Irrespective of the technological means used to process personal data.





# Guidelines for Collecting and Using Speech Data

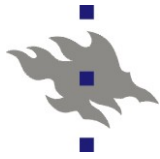
- Can be collected if there is a legitimate purpose.
- On the base of the ***unambiguously given consent***.
  - Who can give?
  - Is informing enough?
  - Notice of the reason for the proceeding
    - Data shall not be further processed in a way incompatible with the purposes for which it was collected.
    - Possible later needs not necessary known.
  - Can be cancelled any time!
- Limited collection
  - Only personal data that is relevant.



# Guidelines for Collecting and Using Personal Data

## Anonymising the personal data

- Anonymised data is not subject to the data protection legislation
- Has to be impossible to identify the individuals from that information and any other information
- Deanonymising?
- If you can anonymise the data –you should do it!



# Guidelines for Collecting and Using Speech Data

## ■ Paradigm shift: Computational Social Science

- Tempting area with lot of problems

- How to get the material?

- Sharing between data collectors and research problematic

- Legislation might limit the possibility to get it from the collectors.
- Tradition can also limit.

- Living Lab/User Communities

- Usually communities are not legal entities, cannot make agreements.
- How to get the permission from the data subject?



# Guidelines for Collecting and Using Speech Data

- Openness!
- Pay attention to form of the consent.
- Follow good data management practices.
  - Secure process of Data (technically and organisationally).
- Data protection offices are there to help!
  - According to Directive 95/46/EC, each member state shall establish in its territory at least one data protection authority.
  - List of data protection authorities:  
[http://ec.europa.eu/justice\\_home/fsj/privacy/nationalcomm/index\\_en.htm](http://ec.europa.eu/justice_home/fsj/privacy/nationalcomm/index_en.htm)



## Exception in Legislation – Scientific Research

- Member States may, where there is clearly no risk of breaching the privacy of the data subject, restrict by a legislative measure the data subjects rights to access to data:
- when data are processed solely for purposes of scientific research, or
- are kept in personal form for a period which does not exceed the period necessary for the sole purpose of creating statistics.

(Data Protection Directive, Article 13)



## Exception in Legislation – Scientific Research

When the data have not been obtained straight from the data subject:

- Data subject has to typically be informed for the first disclosing.
- However, this shall not apply where, in particular for processing for statistical purposes or for the purposes of historical or scientific research, the provision of such information proves impossible or would involve a disproportionate effort.
  - Number of data subjects.
  - The age of the data.
- In these cases Member States shall provide appropriate safeguards.
  - Typically permissions from the public authority.

(Data Directive 95/46/EC, Article 11)