

Dealing with sensitive data

Open Science Summer School

Pablo Diaz, Marieke Heers, FORS

Program of the day

09:15-10:00 Legal framework (personal data)

Break

10:15-11:00 Anonymization

Break

11:15-12:00 Consent and sharing

FORS – Swiss Center of Expertise in the Social Sciences

- **Research infrastructure** of national scope intended for any institution or person active in the social sciences
- **Funded by the Confederation and the Swiss National Science Foundation (SNSF)**, hosted by the University of Lausanne
- **Founded in 2008** by merging existing entities (SIDOS, SHP, Selects)

Main activities

Infrastructure

Project catalogue, data archiving and dissemination service, support

Data production

National (SHP, Selects) and international surveys (ESS, MOSAiCH-ISSP, Share, EVS)

Research

Thematic, methodological and epistemological research

Collaboration

At the national and international level, in all our active domains

Development

Cutting-edge tools, data collection, and linking

Services

Archiving service

Preservation of Swiss social science research data

Data distribution

Access to archived data at FORS

Support

Advice and consultation in our areas of expertise, trainings, workshops, FORS Guides

Survey mandates

Specific surveys or parts of surveys carried out at the request of individuals or institutions

Information

Overview of social science research in Switzerland

Archiving service

SWISS  base

Data archiving

New requirements

Long-term preservation

Enhance the value of
research projects

Data access

Direct access to:

+ 600 data sets

+ 11'000 project
descriptions

Data management

Training

Consultancy

Development of materials
(i.e. guides)

Data protection

Some legal considerations

Pablo Diaz, FORS/UNIL
Open Science Summer School, UNIZH, July 2022

Managing sensitive data: a legal matter ?

Personal and sensitive data cannot be managed without knowledge of existing laws.

- The processing of this type of data is governed by a certain number of rules about **data protection**.
- While data management is never a purely normative issue, the implementation of adequate strategies necessarily requires an understanding of the regulatory framework which provides the fundamental **definitions**, the **limits** not to be exceeded and the **rights** of each person involved.

What is Data Protection ?

Data protection is commonly defined as the **rules and strategies** designed to protect people's **privacy**.

- **Privacy** is someone's right to keep their personal matters and relationships secret (right to express oneself selectively).
- In the field of data protection, privacy is generally referred to through the concept of **informational self-determination** (autonomy).

Data protection and research

- As social scientists, our work often requires the processing of personal data.
- This said, many of the tools (and sometimes methods) we use to collect and process data from our research participants have privacy concerns:
 - Smartphones (cloud backup, geolocation, etc.)
 - Cloud (Dropbox, etc.)
 - Covert observation
- We have to be very careful !

Legal bases

Legal bases

Informational self-determination is guaranteed by a number of **fundamental texts**:

- Universal Declaration of **Human Rights** (art. 12)
- European Convention on **Human Rights** (art. 8)

THE UNIVERSAL DECLARATION OF HUMAN RIGHTS



Adopted by the General Assembly of the United Nations in 1948, the Universal Declaration states fundamental rights and freedoms to which all human beings are entitled.

You have the responsibility to respect the rights of others.

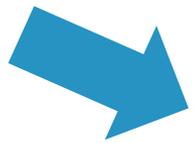
We are all born free and equal.

Everyone is entitled to these rights no matter your race, religion, sex, language, or nationality.

Everyone has the right to life, freedom, and safety.

No one can take away any of your rights.

<p>No one has the right to hold you in slavery.</p> 	<p>You have the right to seek asylum in another country if you are persecuted in your own.</p> 	<p>Every adult has the right to a job, a fair wage, and membership in a trade union.</p> 
<p>No one has the right to torture you.</p> 	<p>Everyone has the right to a nationality.</p> 	<p>You have the right to leisure and rest from work.</p> 
<p>You have a right to be recognized everywhere as a person before the law.</p> 	<p>All consenting adults have the right to marry and to raise a family.</p> 	<p>Everyone has the right to an adequate standard of living for themselves and their family.</p> 
<p>We are all equal before the law and are entitled to equal protection of the law.</p> 	<p>You have the right to own property.</p> 	<p>Everyone has the right to an education.</p> 
<p>You have the right to seek legal help if your rights are violated.</p> 	<p>Everyone has the right to belong to a religion.</p> 	<p>Everyone has the right to freely participate in the culture and scientific advancement of their community, and their intellectual property as artist or scientist should be protected.</p> 
<p>No one has the right to wrongly imprison you or force you to leave your country.</p> 	<p>You have the right to think and voice your opinions freely.</p> 	<p>We are all entitled to a social order in which we may enjoy these rights.</p> 
<p>You have a right to a fair, public trial.</p> 	<p>Everyone has the right to gather as a peaceful assembly.</p> 	<p>Everyone's rights and freedoms should be protected unless they obstruct the rights and freedoms of others.</p> 
<p>Everyone is innocent until proven guilty.</p> 	<p>You have the right to participate in the governance of your country, either directly or by helping to choose representatives in free and genuine elections.</p> 	<p>No State, group, or person can use this Declaration to deny the rights and freedoms of others.</p> 
<p>You have the right to privacy. No one can interfere with your reputation, family, home, or correspondence.</p> 	<p>You have the right to social security and are entitled to economic, social, and cultural help from your government.</p> 	
<p>You have the right to travel.</p> 		



Legal bases

Swiss Federal **Constitution:**

« Every person has the right to privacy in their private and family life and in their home, and in relation to their mail and telecommunications. » (art. 13 al. 1)

« Every person has the right to be protected against the misuse of their personal data « (art. 13 al. 2)

Legal bases

In Switzerland, there are laws at two levels:

- Federal (**Constitution, FADP**)
- Cantonal (**LPrD**, etc.)

In Europe there is a General Data Protection Regulation (**GDPR**)

At the international level 128 countries out of 194 have data protection legislation:

https://unctad.org/en/Pages/DTL/STI_and_ICTs/ICT4D-Legislation/eCom-Data-Protection-Laws.aspx

Legal bases

- In order to know whether a data protection law applies, it is first necessary to know whether or not **personal data** is being processed.
- If there is no personal data, the data protection laws do not apply.
- It is therefore essential to learn how to recognize personal data (as well as sensitive data, which is a subcategory of personal data)

PERSONAL DATA

Am I processing personal data ?

Processing

Any operation with data, irrespective of the means applied and the procedure, and in particular the collection, storage, use, revision, disclosure, archiving or destruction of data. (Art. 3 lit. a FADP)

Personal data

All information relating to an identified person.
(Art. 3 lit. a FADP)

Sensitive data

Personal data on:

1. Religious, ideological, political or trade-union; related views or activities
 2. Health, the intimate sphere or the racial origin;
 3. Social security measures;
 4. Administrative or criminal proceedings and sanctions.
- (Art 3. lit. c FADP)

Personal data

“All information relating to an identified person” (Art. 3 lit. a FADP)

Very broad notion: everything that can be related to a specific person is personal data !

The most common: name, date of birth, home address, phone number, email, IP address, picture, etc.

But also: opinions, original ideas, a style of writing, the way of walking, etc.

Personal data can be “objective” or “subjective”

Examples of personal data

Contact details

First name: Paul
Last name: Dupont
Phone number: 123456
Email: paul@dupont.ch

Quotes

“I would like to die on Mars. Just not on impact”

Picture



The way you dance is unique, and computers can tell it's you

Nearly everyone responds to music with movement, whether through subtle toe-tapping or an all-out boogie. A recent discovery shows that our dance style is almost always the same, regardless of the type of music, and a computer can identify the dancer with astounding accuracy.



Studying how people move to music is a powerful tool for researchers looking to understand how and why music affects us the way it does. Over the last few years, researchers at the Centre for Interdisciplinary Music Research at the University of Jyväskylä in Finland have used motion capture technology—the same kind used in Hollywood—to learn that your dance moves say a lot about you, such as how extroverted or neurotic you are, what mood you happen to be in, and even how much you empathize with other people.

Chinese 'gait recognition' tech IDs people by how they walk

By DAKE KANG November 6, 2018



Click to copy

BEIJING (AP) — Chinese authorities have begun deploying a new surveillance tool: “gait recognition” software that uses people’s body shapes and how they walk to identify them, even when their faces are hidden from cameras.

Even Anonymous Coders Leave Fingerprints

Researchers have repeatedly shown that writing samples, even those in artificial languages, contain a unique fingerprint that's hard to hide.

RESEARCHERS WHO STUDY stylometry—the statistical analysis of linguistic style—have long known that writing is a unique, individualistic process. The vocabulary you select, your syntax, and your grammatical decisions leave behind a signature. Automated tools can now accurately identify the author of a forum post for example, as long as they have adequate training data to work with. But newer research shows that stylometry can also apply to *artificial* language samples, like code. Software developers, it turns out, leave behind a fingerprint as well.

Rachel Greenstadt, an associate professor of computer science at Drexel University, and Aylin Caliskan, Greenstadt's former PhD student and now an assistant professor at George Washington University, have found that code, like other forms of stylistic expression, are not anonymous. At the DefCon hacking conference Friday, the pair will present a number of studies they've conducted using machine learning techniques to de-anonymize the authors of code samples. Their work, some of which was funded by and conducted in collaboration with the United States Army Research Laboratory, could be useful in a plagiarism dispute, for instance, but also has privacy implications, especially for the thousands of developers who contribute open source code to the world.



**(DON'T) WORRY YOU ARE
UNIQUE!**

Sensitive data

Personal data on: religious, ideological, political or trade-union related views or activities; health, the intimate sphere or the racial origin; social security measures; administrative or criminal proceedings and sanctions (Art3. lit. c FADP)

The list provided by FADP is **exhaustive** (e.g. in Switzerland salary is not considered sensitive data)

That said, depending on the **context**, almost all data can be considered sensitive (name, photo, job, etc.)

Examples of sensitive data

Contact details

First name: Pedro
Last name: Ruiz
Gender: Trans
Job: Trade-Unionist

Quote

“We are now in the process of defeating the radical left, the Marxists, the anarchists, the agitators, the looters, and people who, in many instances, have absolutely no clue what they are doing”

Picture



Sensitive data ?



Sensitive data ?



Kayan people (Myanmar)

From Wikipedia, the free encyclopedia

(Redirected from [Kayan \(Burma\)](#))

For the ethnic group from [Borneo](#), see [Kayan people \(Borneo\)](#).

The **Kayan** are a sub-group of Red Karen ([Karenni people](#)), [Tibeto-Burman ethnic minority](#) of [Myanmar](#) (Burma). The Kayan consists of the following groups: Kayan Lahwi (also called **Padaung**, ပဒေါင်း[bədàʊɰ]), Kayan Ka Khaung (Gekho), Kayan Lahta, Kayan Ka Ngan. Kayan Gebar, Kayan Kakhi and, sometimes, [Bwe people \(Kayaw\)](#). They are distinct from, and not to be confused with, the [Kayan people of Borneo](#).

Padaung (Yan Pa DOUNG) is a [Shan](#) term for the Kayan Lahwi (the group in which women wear the brass [neck rings](#)). The Kayan residents in Mae Hong Son Province in Northern Thailand refer to themselves as Kayan and object to being called Padaung. In *The Hardy Padaungs* (1967) [Khin Maung Nyunt](#), one of the first authors to use the term "Kayan", says that the Padaung prefer to be called Kayan.^[1] On the other hand, [Pascal Khoo Thwe](#) calls his people Padaung in his 2002 memoir, *From the Land of Green Ghosts: A Burmese Odyssey*.^[2]

In the late 1980s and early 1990s due to conflict with the military regime in Myanmar, many Kayan tribes fled to the [Thai](#) border area.^[3] Among the refugee camps set up there was a Long Neck section, which became a tourist site, self-sufficient on tourist revenue and not needing financial assistance.^[4]

Kayan

ကယန်း



Sensitive data ?



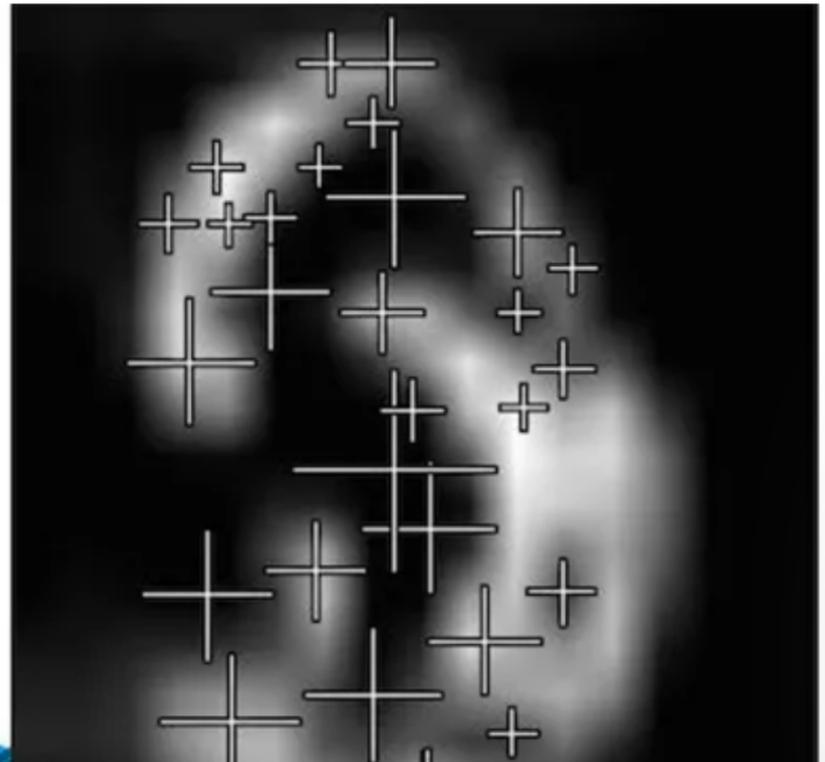
Bodyprint could let you unlock your phone with your ear print

An experimental system uses a smartphone's capacitive screen as a low-res sensor keyed to the user's body prints.



Michelle Starr  April 27, 2015 8:22 p.m. PT

ES



The importance of databases / knowledge

Of course, for certain information to be linked to someone, a number of conditions must be met:

- There must be the existence of databases that link the information to an identity (e.g. without databases, biometric information is “useless”).
- The Internet can function as a giant database (a lot of data is available via google).
- It is sometimes necessary to know the persons to recognize them (importance of who has access to the information)

Is public data personal data ?

- Sometimes people make their personal data available to everyone. For example via social networks, blogs, websites, the press, etc.
- Even if this data could be qualified as “public”, it is still personal data and must be treated with care.
- It is important to check the privacy policies of the places where data is collected. Some types of processing may be prohibited.

Am I processing personal data?

Yes  **DP Laws apply**

No  **DP Laws don't apply**

 If you plan to anonymize the data, be aware that they remain personal until you do so (DP laws apply until anonymization is complete).

Am I processing personal data?

- In the social sciences, it is very difficult to have completely anonymous data.
- To be considered anonymous, **all information** that can be linked to an identifiable person must be **permanently destroyed**.
- However, as we have seen, a lot of information can potentially allow the identification of an individual.
- It is therefore generally **safer to assume that we are dealing with personal data**.

Data protection

General principles

Main legal principles in Switzerland

Swiss law sets out 6 main principles of data protection that must always be respected (cumulatively):

1. Personal data may only be processed **lawfully**
2. Its processing must be carried out in **good faith** and in a **transparent** manner
3. Data collection must be **proportionate**
4. Personal data may only be processed for the **purpose** indicated at the time of collection
5. Anyone who processes personal data must make certain that it is **correct**
6. Personal data must be protected against unauthorised processing through adequate **technical and organisational measures**.

Main legal principles in Switzerland

P.1. Personal data may only be processed **lawfully**

- The general rule is normally that data processing must not be prohibited.
- However, when working for a public body (university, etc.), the rule is stricter: data processing **must be legally permitted**.
- This means that you must have a **legal basis** for processing personal data!

 Consent is often the only legal basis available

Main legal principles in Switzerland

P.2. Personal data must be processed in **good faith** and in a **transparent** manner

- Individuals must be **informed** about the processing of their data
- Individuals must be able to acknowledge that their data is being collected
- People should not be deceived

Main legal principles in Switzerland

P.3. Data collection must be **proportionate**

- Only collect the data you need
- You can't collect data just “because it might be useful.”
- The necessary data cannot always be determined in advance. In this case it is necessary to be in good faith and to justify it.

Main legal principles in Switzerland

P.4. Personal data may only be processed for the **purpose** indicated at the time of collection

- Normally, data can only be used for the purpose originally intended (and communicated to individuals).
- There is, however, an exception for research:
 - Data collected for a non-research purpose may be used for research purposes (e.g. government data).
 - ⚠ Data collected for research purposes cannot be used for other purposes.

Main legal principles in Switzerland

P.5. Anyone who processes personal data must make certain that it is **correct**

- The information processed must be correct.
- If it is not, it must be corrected.

Main legal principles in Switzerland

P.6. Personal data must be protected against unauthorised processing through adequate **technical and organisational measures.**

- Encryption
- Secure storage
- Access management
- Etc.

Thank you!