

An Introduction to Trustworthy Machine Learning

Short course

23.11.2022

Lecturers

- Daniel Gatica-Perez
 - Head of Social Computing Group at Idiap Research Institute and Professor at EPFL
 - Research interests: Social and Ubiquitous Computing
- Sina Sajadmanesh
 - Research Assistant at Idiap Research Institute and PhD student at EPFL
 - Research interests: Data Privacy, Trustworthy ML, Graph Neural Nets
- Ali Shahin Shamsabadi
 - Research Associate at The Alan Turing Institute, AI Programme
 - Research interests: Data Privacy and Trustworthy ML



Course Content

- Day 1
 - Introduction to privacy and personal data
 - Introduction to differential privacy
 - Differentially private machine learning
 - Hands-on exercises
- Day 2
 - Introduction to adversarial examples
 - Defenses against adversarial examples
 - Adversarial examples for privacy protection
 - Hands-on exercises

Introduction to privacy and personal data

Daniel Gatica-Perez

this lecture

defining privacy

case study: facebook and the real-name web

regulations on personal data





2018
credit: Time Magazine

the multifaceted nature of privacy

law

sociology

computing

psychology

“Privacy and technology are closely intertwined. Shifts in technology requires us to rethink our attitude toward privacy”



defining privacy

Privacy as “the right to be let alone” (Warren & Brandeis, 1890)

Information privacy: “The claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others” (Westin, 1967)

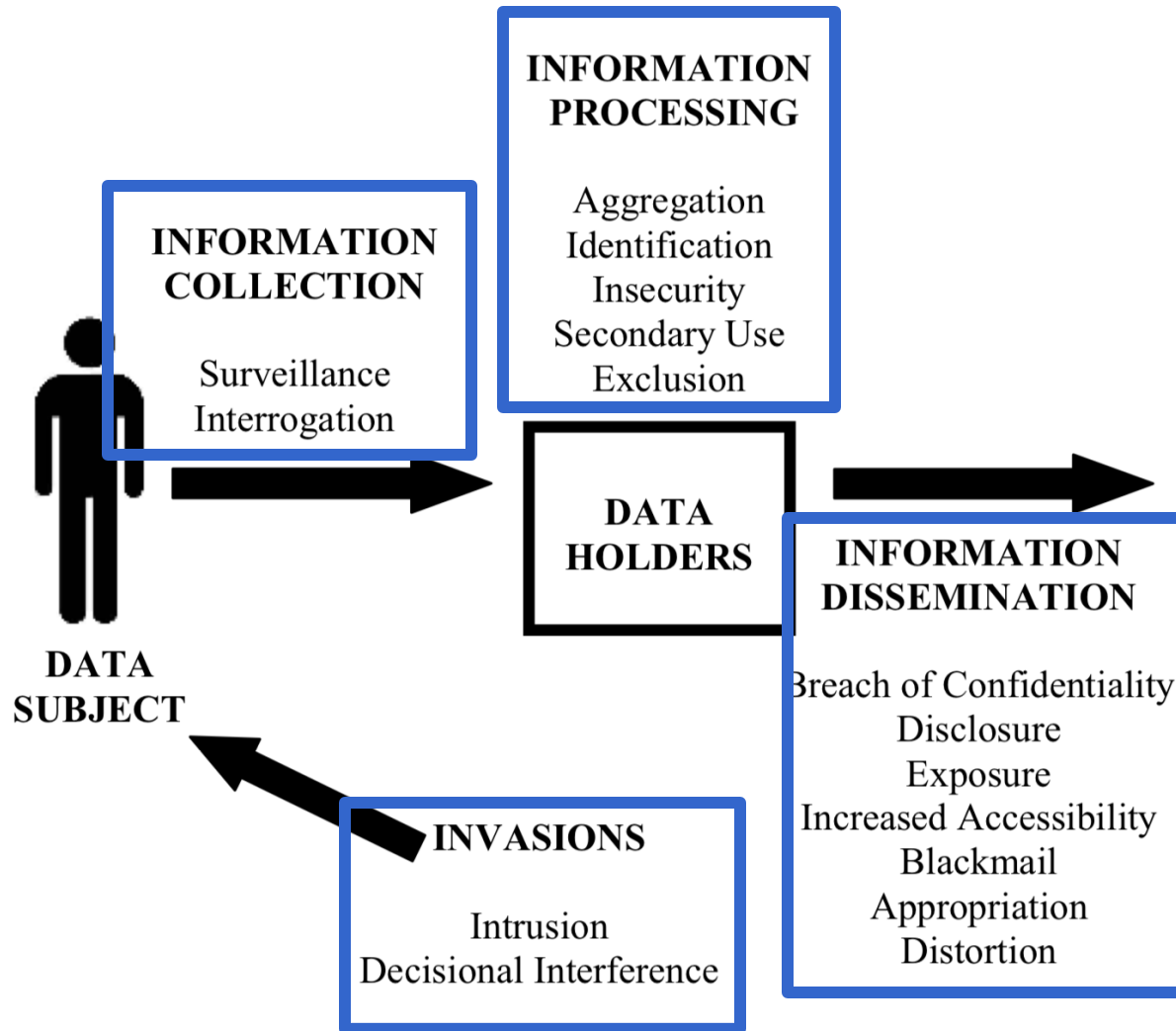
“A final definition of privacy is difficult. It is related, but not identical with, secrecy, solitude, liberty, autonomy, freedom, and intimacy”

“Privacy is often not a goal in itself, ..., but rather an expectation of being in a state of protection without having to actively pursue it”

Photo credit: Todd Diemer on Unsplash

Solove's privacy taxonomy (2006)

activities that might lead to privacy problems



Credit: D. J. Solove, A Taxonomy of Privacy, University of Pennsylvania Law Review, Vol. 154, No. 3, Jan. 2006

M. Langheinrich, Privacy in Ubiquitous Computing. In J. Krumm, (Ed.), Ubiquitous Computing Fundamentals, CRC Press, 2010

defining personal data



EU General Data Protection Regulation (GDPR), Article 4:

” ‘Personal data’ means any information relating to an identified or identifiable natural person (‘data subject’); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person”

this lecture

defining privacy

case study: facebook and the real-name web

regulations on personal data

Facebook's Name Policy

▼ What names are allowed on Facebook?

Personal Accounts

Facebook is a community where people use their real identities. We require everyone to provide their **real names**, so you always know who you're connecting with. This helps keep our community safe.

Names can't include:

- Symbols, numbers, unusual capitalization, repeating characters or punctuation
- Characters from multiple languages
- Titles of any kind (ex: professional, religious, etc)
- Words, phrases, or nicknames in place of a middle name
- Offensive or suggestive content of any kind

Other things to keep in mind:

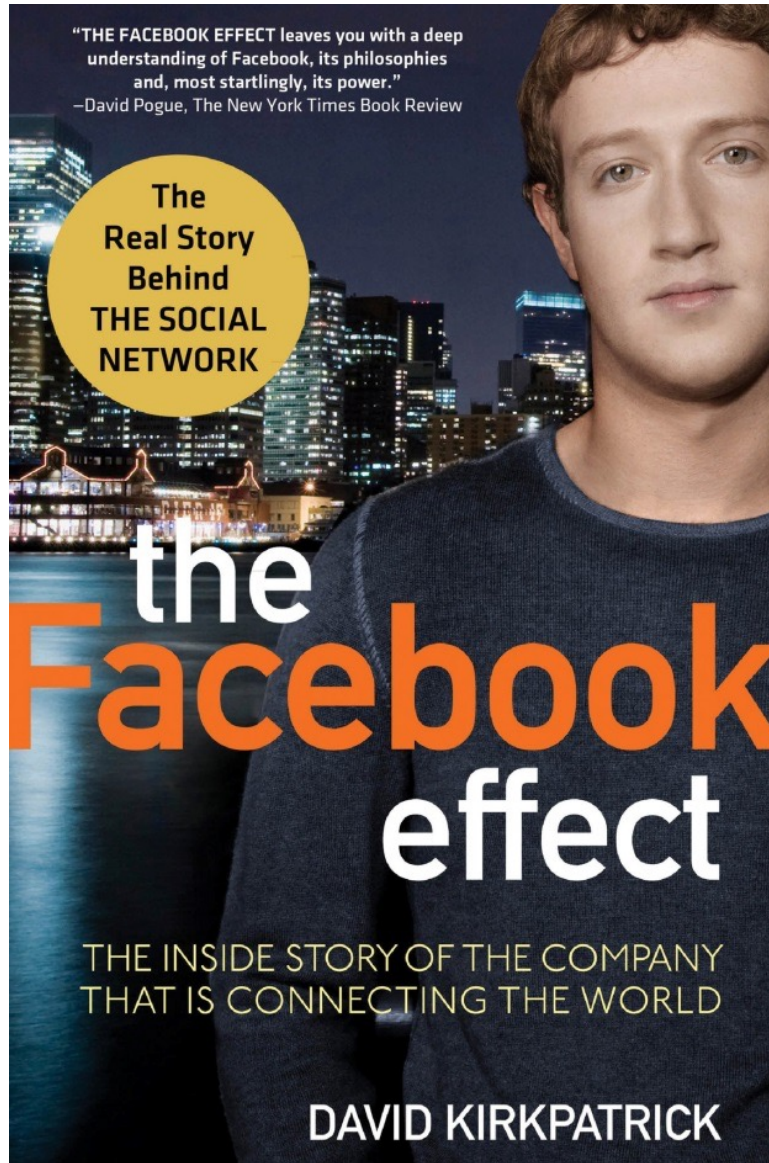
- The name you use should be your real name as it would be listed on your credit card, student ID, etc.
- Nicknames can be used as a first or middle name if they're a variation of your real first or last name (like Bob instead of Robert)
- You can also list another name on your account (ex: maiden name, nickname, or professional name), by adding an [alternate name](#) to your Timeline
- Only one person's name should be listed on the account – Timelines are for individual use only
- Pretending to be anything or anyone is not allowed

Original link (no longer available):

<https://www.facebook.com/help/292517374180078>

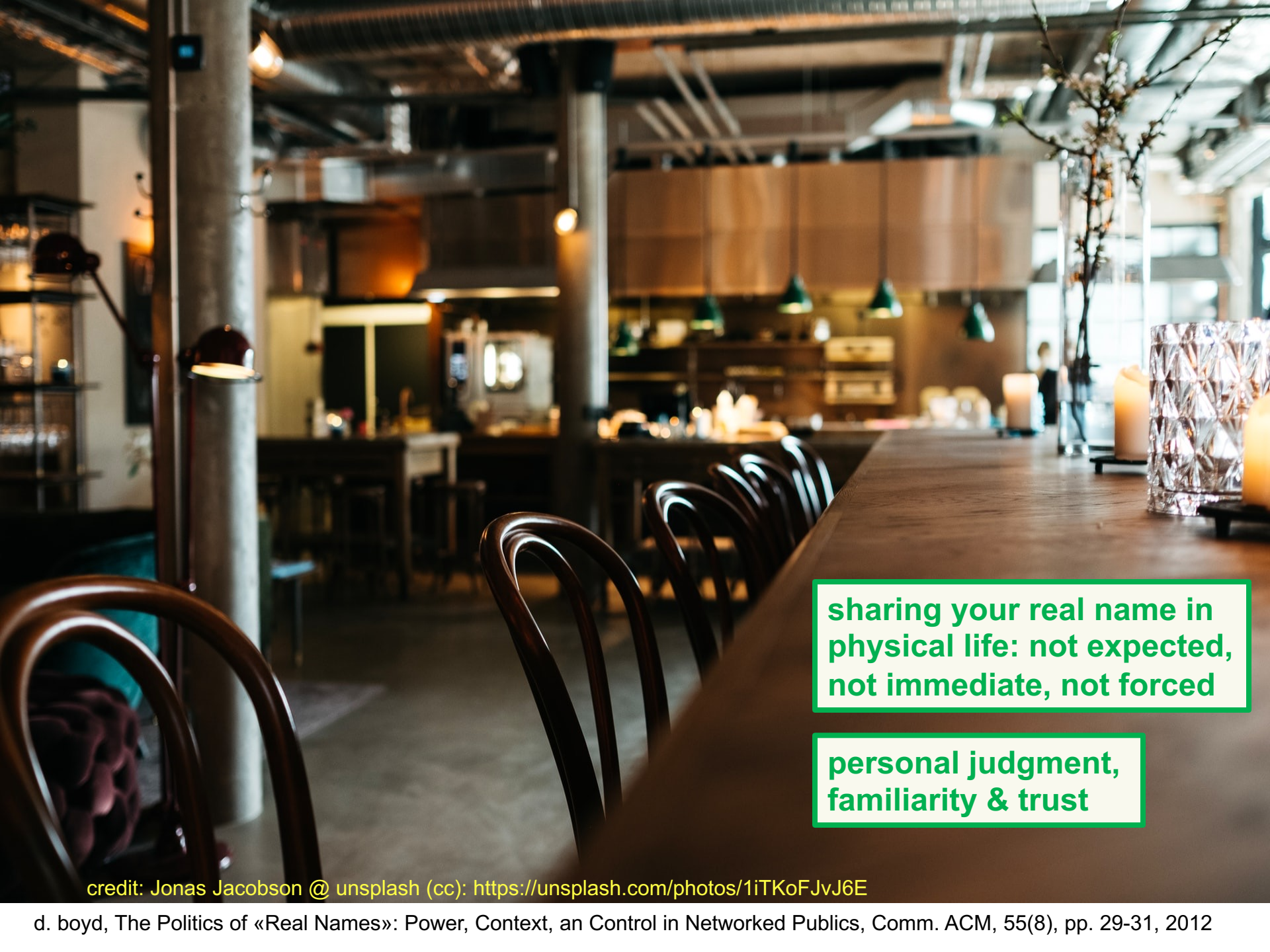
Current version (accessed Nov 2022):

<https://www.facebook.com/help/112146705538576/>



“Having two identities for yourself is an example of a lack of integrity”

«the real-name web is not a technology,
it is a social practice and a system of values»



sharing your real name in
physical life: not expected,
not immediate, not forced

personal judgment,
familiarity & trust

credit: Jonas Jacobson @ unsplash (cc): <https://unsplash.com/photos/1iTKoFJvJ6E>

the pre-2.0 web was

+ textual & simplified

- + no images, audio, video
- + users were authors of text
- + easier to be someone else

+ sparsely connected

- + discussions created about topics & interests, not people

+ strange

- + biased to tech-skilled people
- + one never knew exactly who was on the other side of screen

credit: OiMax @flickr (cc):
<https://www.flickr.com/photos/oimax/2141263830>

the real-name web is

+ detailed

- + images, audio, video
- + difficult to be someone else

+ densely connected

- + friends & family are online
- + discussions about people

+ familiar & day-to-day

- + the more people, the less strange
- + from “a place out there” to “data about here”

credit: Solen Feyissa @unsplash (cc):
https://unsplash.com/photos/iurEAYyU_c

implications: real-name photo tagging

FB photo tagging (fall 2005)

- + only one way: real names
- + not objects, scenes, topics
as other platforms
- + became world's largest photo site

uses:

- + **access** to personal data and physical appearance
- + **generation** of relational data (events, groups)
- + **identity verification** when device is not identified
- + **labels** for machine learning

this lecture

defining privacy

case study: facebook and the real-name web

regulations on personal data

regulations on personal data

4.5.2016

EN

Official Journal of the European Union

L 119/1

I
(Legislative acts)

REGULATIONS

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)

GDPR: European
General Data
Protection
Regulation



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

The federal Council
The portal of the Swiss government

Federal Council	Federal Presidency	Departments	Federal Chancellery	Federal law	Documentation
▼	▼	▼	▼	▼	▼

[Start](#) > [Federal law](#) > [Classified compilation](#) > [Internal laws](#) > 235.1 Federal Act of 19 June 1992 on Data Protection (FADP)

235.1

[expand all](#) | [article overview](#) | [collapse all](#) |

*English is not an official language of the Swiss Confederation.
This translation is provided for information purposes only and
has no legal force.*

Federal Act on Data Protection

(FADP)

Additional informations

This text is in force.

Decision 19 June 1992

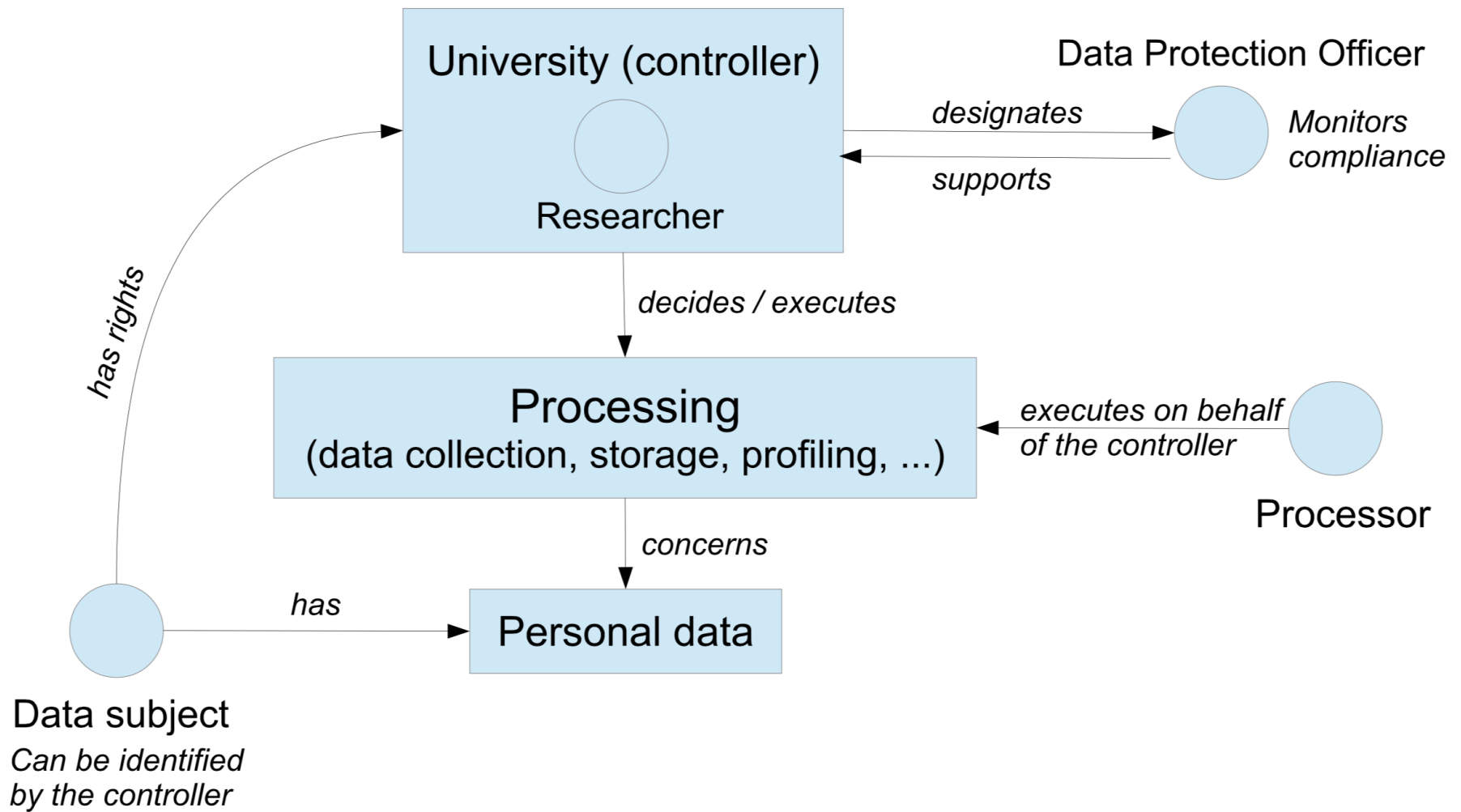
In force 1 July 1993

Tools

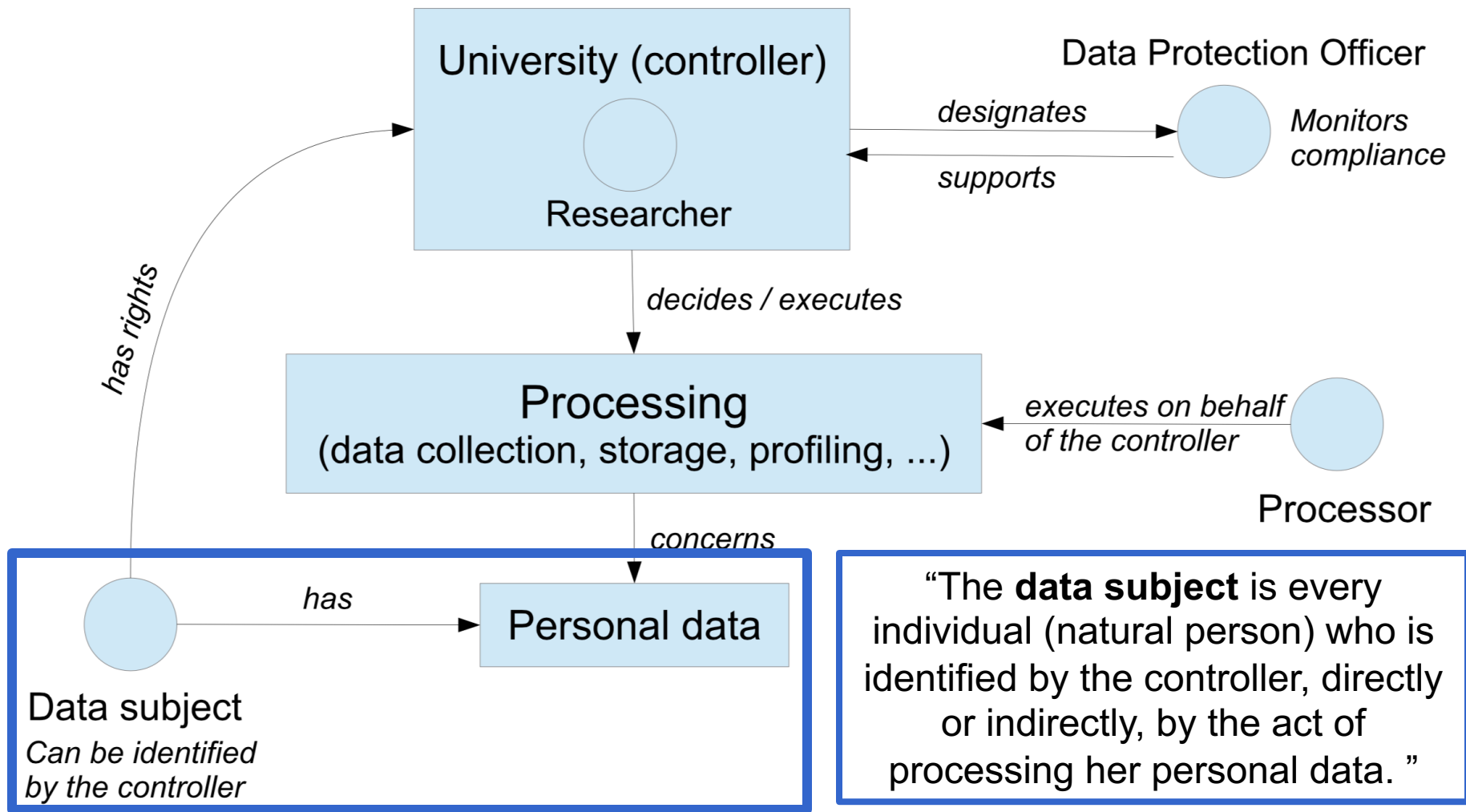
FADP: Swiss Federal
Act of Data Protection

nFADP: new version to
entry into force in Sep
2023

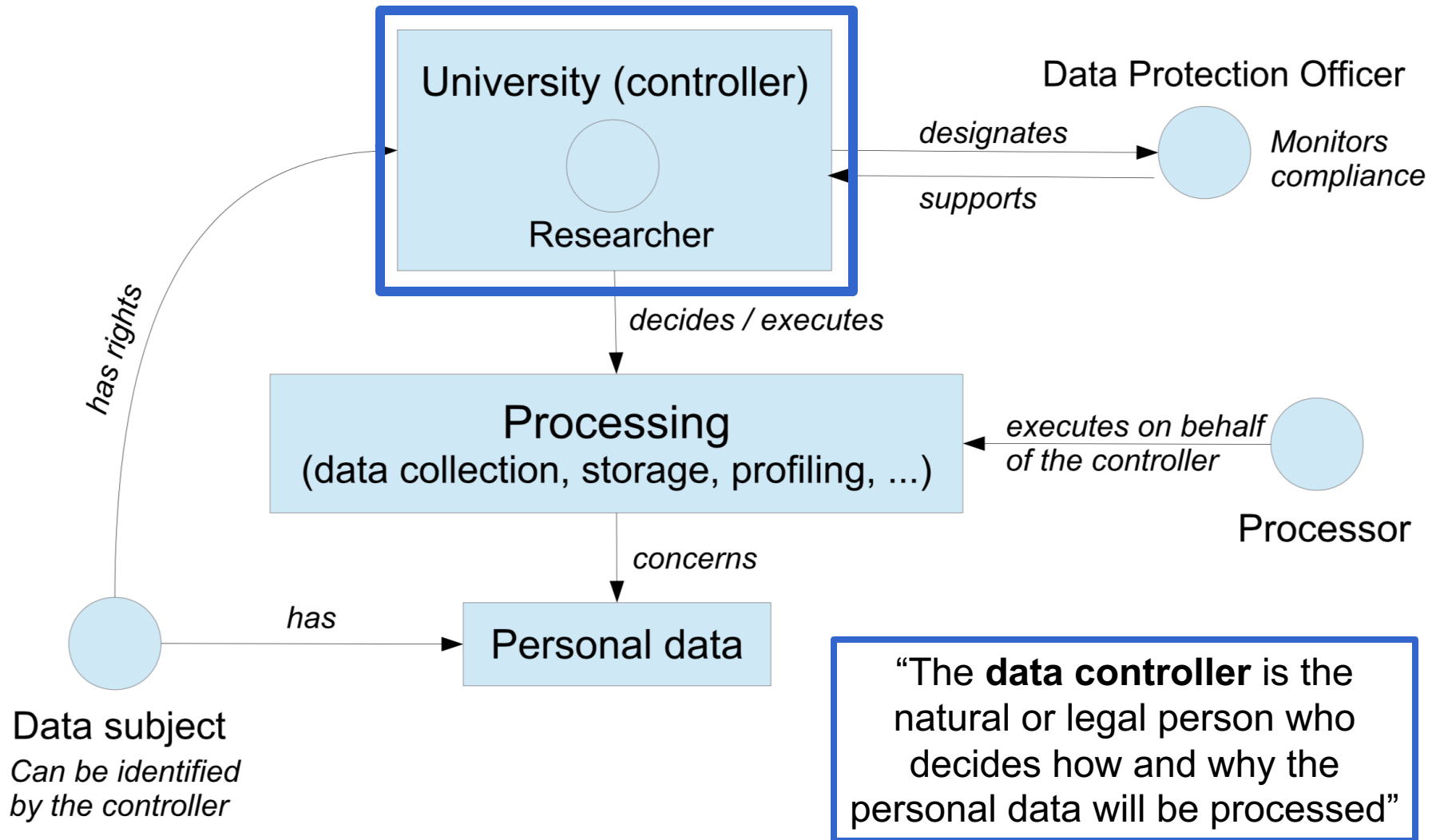
typical configuration of GDPR “ecosystem”



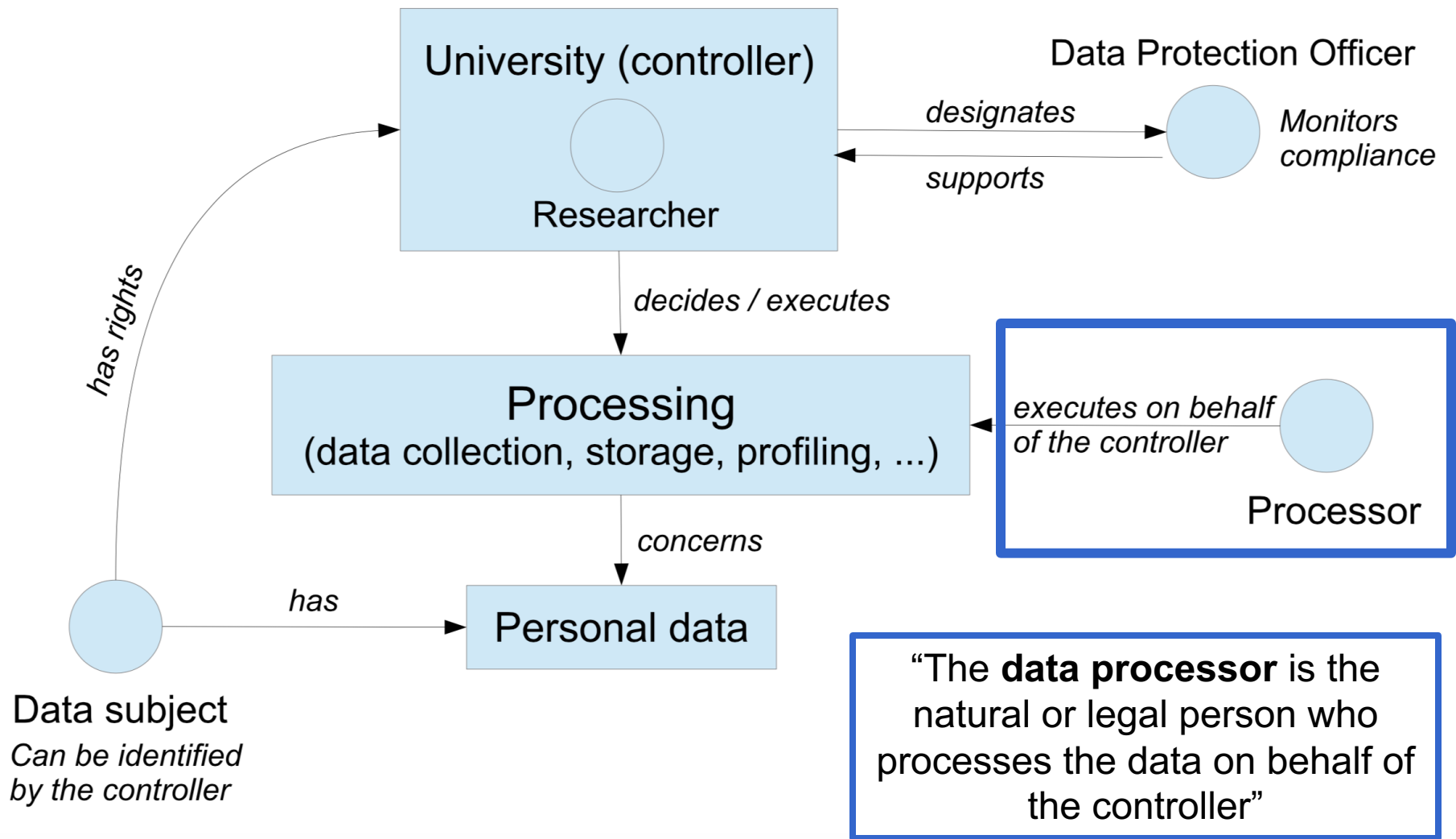
typical configuration of GDPR “ecosystem”



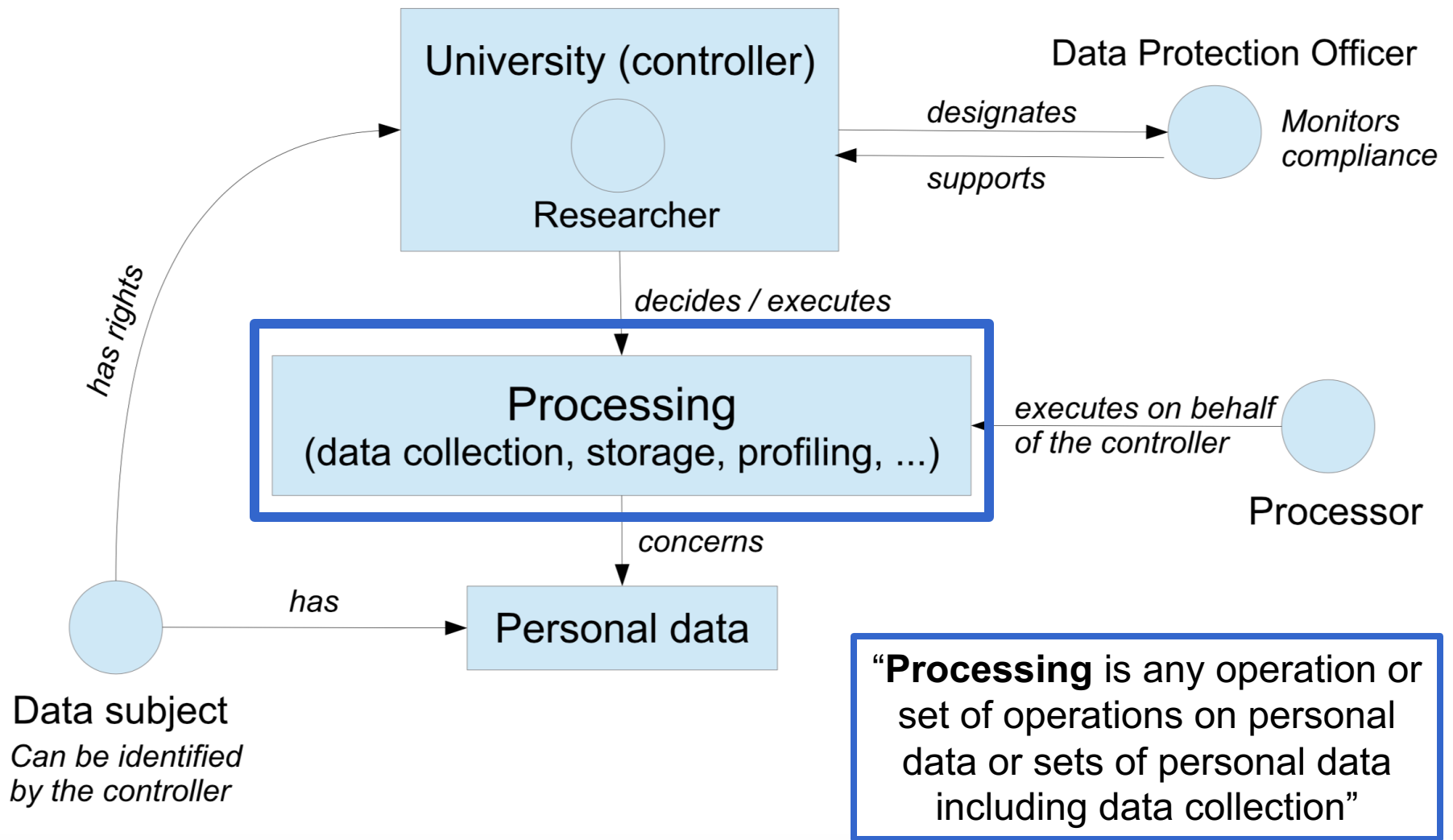
typical configuration of GDPR “ecosystem”



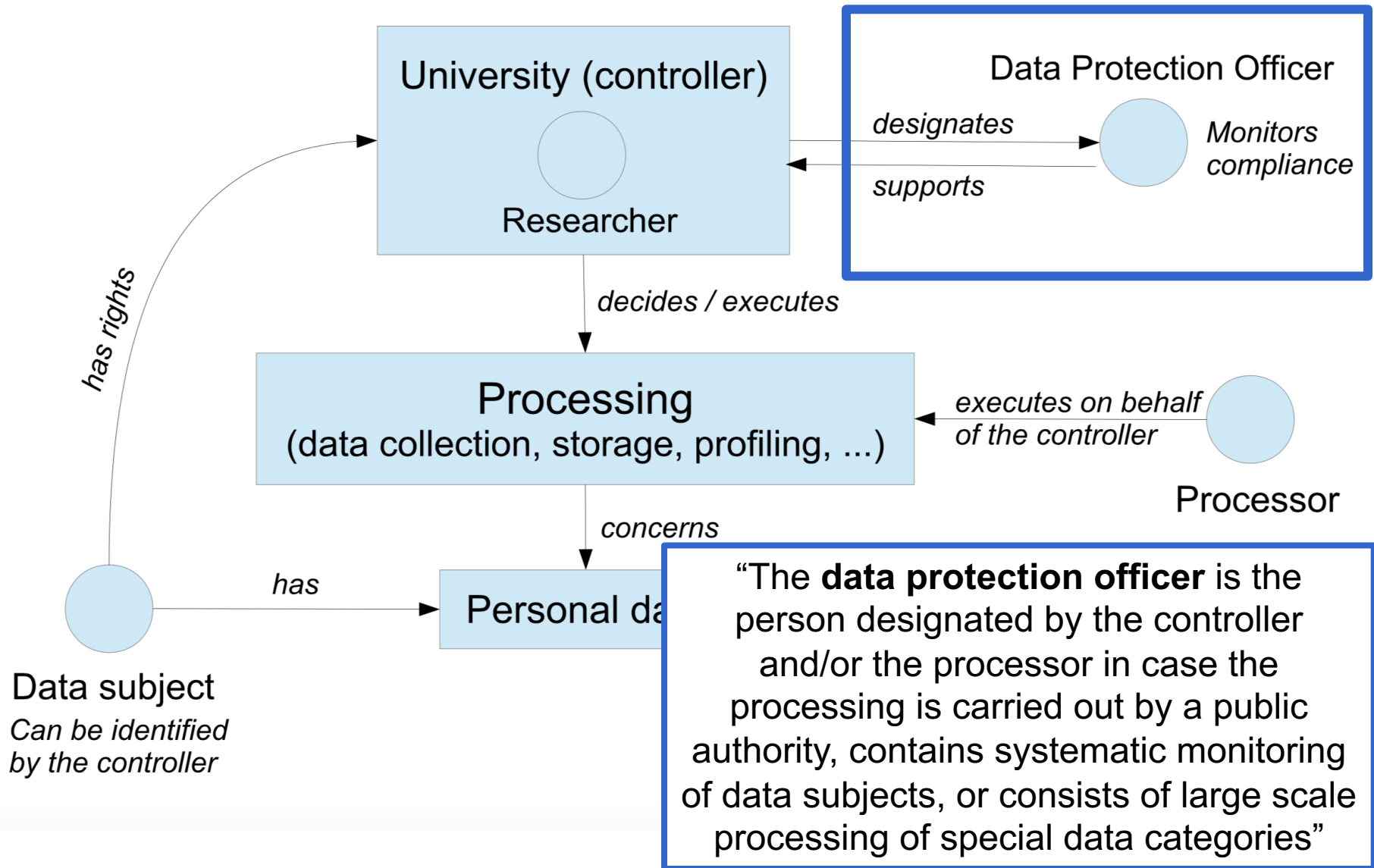
typical configuration of GDPR “ecosystem”



typical configuration of GDPR “ecosystem”



typical configuration of GDPR “ecosystem”



what to remember

privacy

- it is not a new societal issue

- tech has opened up new ways that affect privacy

the real-name web

- it is not a technology, but a practice and a value system

- it enables identity services and machine learning

- it is a source of privacy risks & power disparities

regulations for personal data

- legal frameworks at national & regional levels

- but no worldwide rules

questions?