EPFL

Data Champions initiative at EPFL
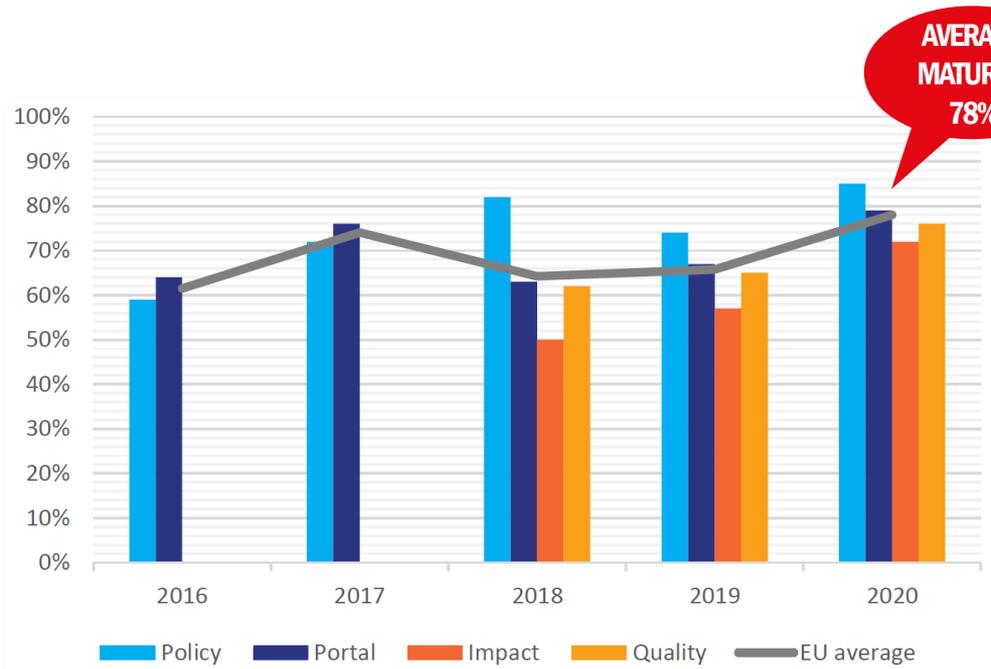
Implementing
FAIR practices

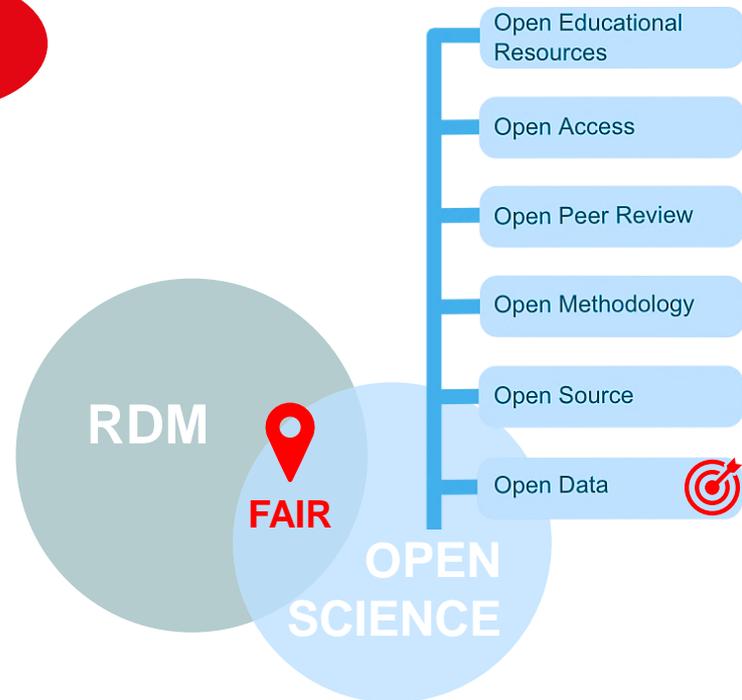**Francesco Varrato**

Research Data
EPFL Library Team

RDM → FAIR → Open Data?
Good practices already exist

EU27 Open Data Maturity Report 2020

AVERAGE MATURITY 78%

Chart legend: Policy, Portal, Impact, Quality, EU average

RDM

FAIR

OPEN SCIENCE

- Open Educational Resources
- Open Access
- Open Peer Review
- Open Methodology
- Open Source
- Open Data

F. Varrato

LIBER2021

Bibliothèque

RDM directions
Top → Down approach

F. Varrato

Institutional data policies

- MIT
- UNIGE
- TU Delft
- University of Oxford
- University of Edinburgh
- University of Cambridge
- Humboldt-Universität zu Berlin
- …

General data guidelines

- SNSF Open Research Data policies
- EC Data Management manual for Horizon 2020
- DLCM Research Data Management Policy Template
- …

@EPFL

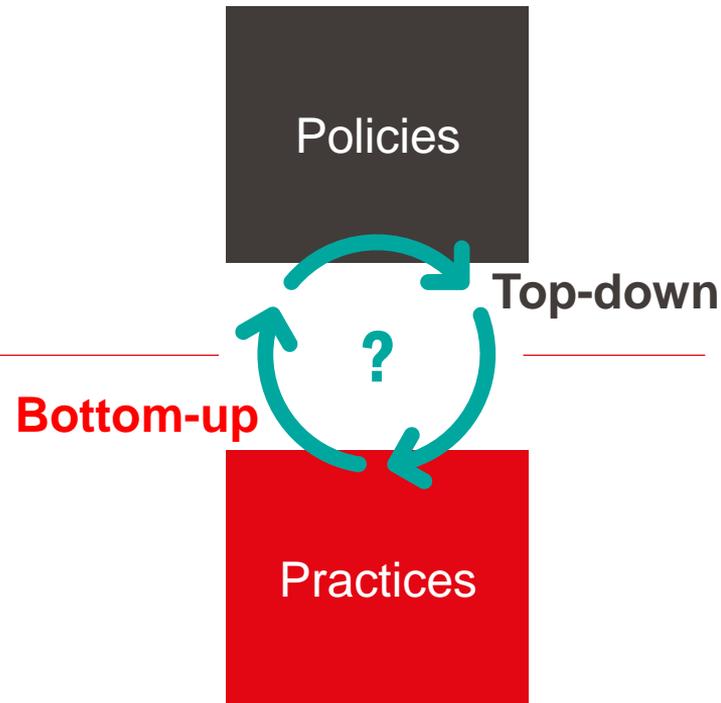- no official Data Policy
- Open Science Scientific Committee

Images: pixabay.com/fr/users/kmicican-2305081

■ Bibliothèque

F. Varrato

EPFL & Cultural shift
No unique approach

STRUCTURE

- 2 **Colleges**
- 5 **Schools**
- > 15 interdisciplinary **centers**
- 26 **Institutes**
- > 370 **laboratories** & research groups

PEOPLE

- > 9,500 **students**
- > 2,200 **PhD** students
- > 3,500 **staff**
- > 350 **professors**

Policies

Top-down

?

Bottom-up

Practices

LIBER2021

Bibliothèque

Data Champions @ EPFL

A possible approach

What / Who

- **Interdisciplinary** community
- Data practices **advocates**
- **Voluntary** researchers & staff

Shared principles

- **FAIR** data principles
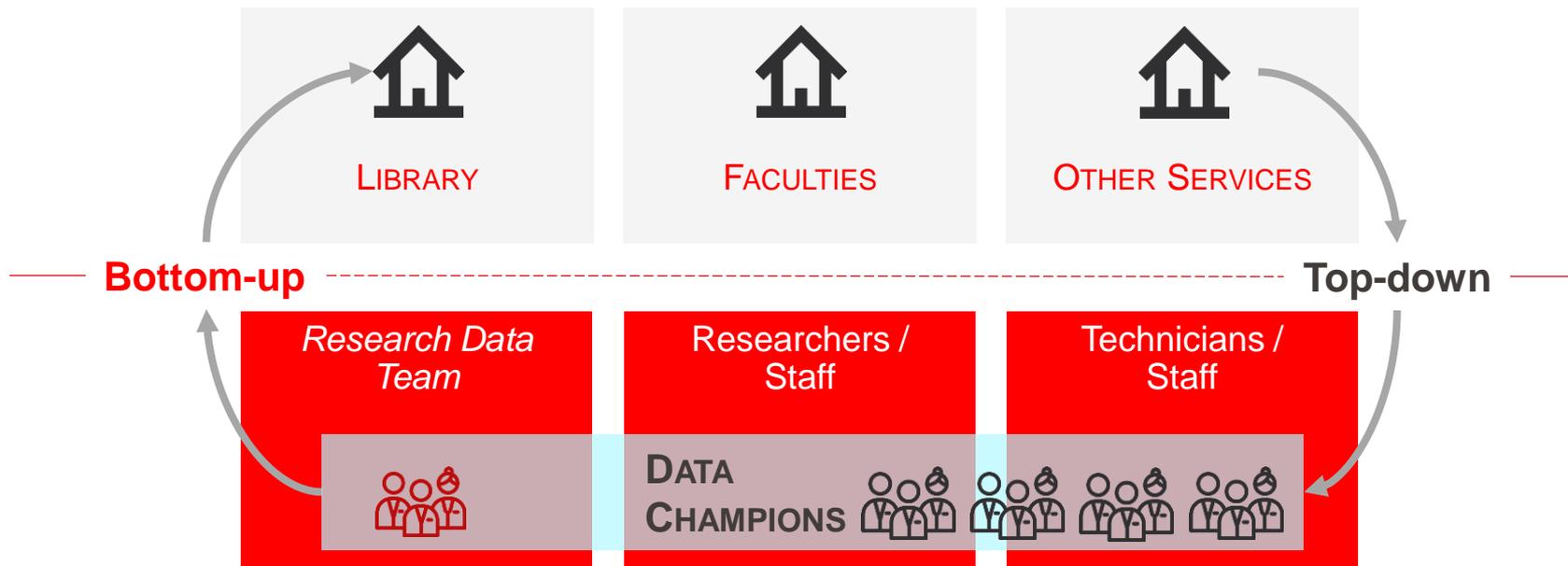- **Passion** for data ♥

Mission(s)

- **Advise** research community
- **Advocate** FAIR data principles
- **Promote** good data mgmt

Platform(s)

- <u>Webpage</u> (<u>Library</u> + <u>VPR</u>)
- Google Drive + <u>GitHub</u>
- Email + Slack + <u>LinkedIn</u>

■ Bibliothèque

F. Varrato

No real barriers
Synergy with central services

| LIBRARY | FACULTIES | OTHER SERVICES |

Bottom-up - **Top-down**

| *Research Data Team* | Researchers / Staff | Technicians / Staff |

DATA CHAMPIONS

- Library provides organization & support
- Library maintains web presence & comm.

LIBER2021

Bibliothèque

Why join the DC community?
Bottom → Up approach

F. Varrato

To **share** my skills and **gain** new ideas

data sharing is essential to combat the reproducibility crisis

I am passionate about rigorous methods and the inductive / abductive process of analyzing **qualitative data**

Improve on data **sorting, traceability, archiving** and **sharing** at EPFL

Wish to help / **encourage** others in the community for a **cultural** shift

Since **similar challenges** can be expected to reoccur in different labs, it seems natural and beneficial to build a community and **share** experience

I am very much committed to **revolutionize** the scientific **community** to become truly collective

I think it is important to **share** information about data management within the School, and **train** each other

I would love to explore **new data** formats, analysis techniques, software, best **practices**

I want to **share** my experiences of using big and **open** data in a discipline that is rather conservative
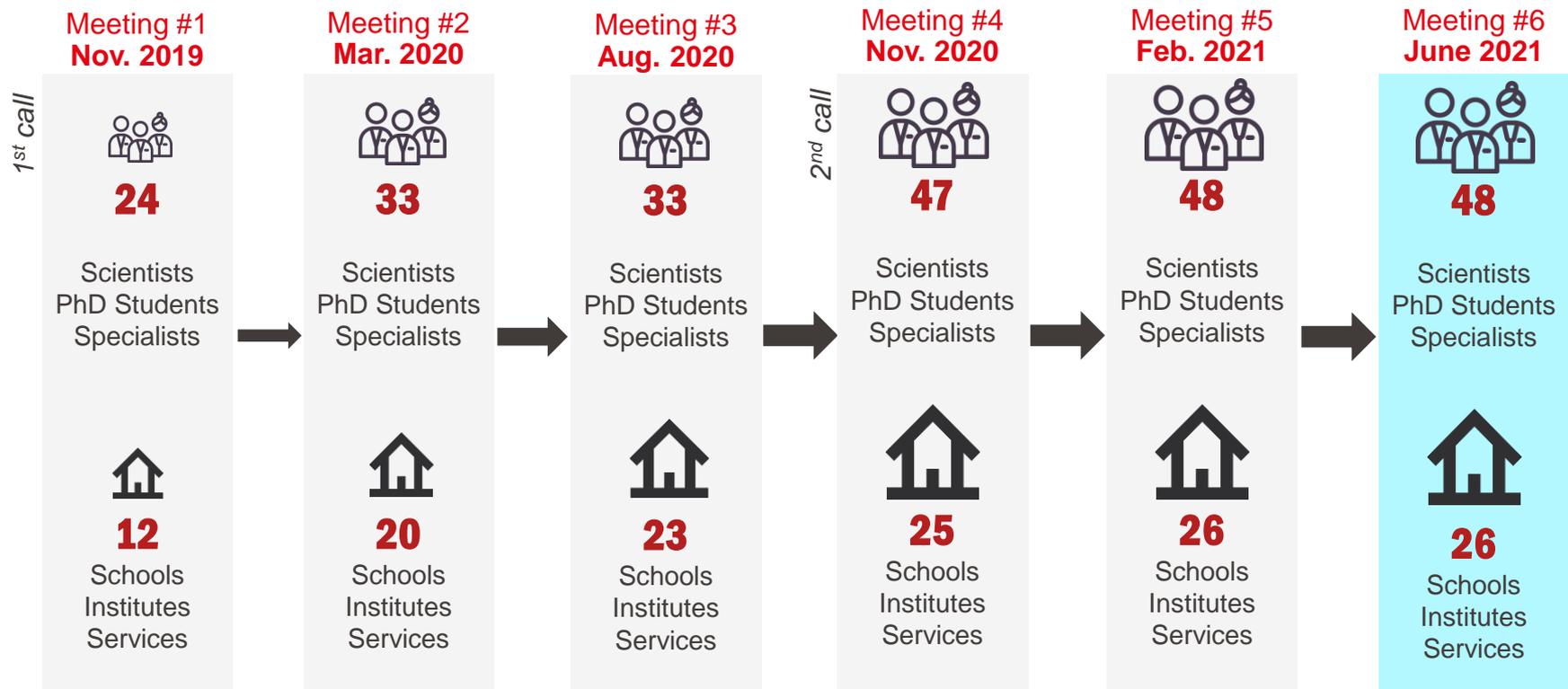
I am big advocate of data **sharing** and effective data visualization as I believe it makes science **better** and more **accessible**

keen to **help** people understanding the **legal** framework

This would be both useful for my everyday **work** in the lab and maybe even my **future**

I would like to **help** fellow researchers with **managing**, acquiring, producing and analyzing their data

LIBER2021

Bibliothèque

EPFL DCs
Some numbers

Meeting #1 **Nov. 2019**	Meeting #2 **Mar. 2020**	Meeting #3 **Aug. 2020**	Meeting #4 **Nov. 2020**	Meeting #5 **Feb. 2021**	Meeting #6 **June 2021**
24	**33**	**33**	**47**	**48**	**48**
Scientists PhD Students Specialists	Scientists PhD Students Specialists	Scientists PhD Students Specialists	Scientists PhD Students Specialists	Scientists PhD Students Specialists	Scientists PhD Students Specialists
12	**20**	**23**	**25**	**26**	**26**
Schools Institutes Services	Schools Institutes Services	Schools Institutes Services	Schools Institutes Services	Schools Institutes Services	Schools Institutes Services

1st call   *2nd call*

+ monthly, more informal, shorter *DCBreaks*

Bibliothèque

Meetings' themes

Sharing expertise, tools, views

Open-Source icons	ELN EPFL	Dataset Peer-Reviewing	Swiss data policies	File naming conventions
Fully reproducible dataset	Art and Dataviz	Open biologic database	Civil Engineering data	Open Data Metrics
EPFL Data Policy	EPFL Data Archive	Qualitative data analysis	Reproducible code platform	Digital Musicology
Research indicators	Preservation *vs.* Publication	Code publishing	Role of Data Manager	Data Protection

→ **FAIR**

Other examples
Sharing expertise, tools, views

The DC community actively collaborated to

- Data Q&A on webpage

- Library's training "*The power of metadata*" + RDM Credited Course for Chem. PhDs

- GitHub repo

- Collaborative documents (Data Policy, EPFL DMP, Data Tools & Platforms, …)

- Help others about

  o Open education platform

  o Publishing reproducible code

  o Python pickle files for data preservation

  o Tips for Data security

  o DMP template for new federal institute

  o …

DC

≠

DATA STEWARD
DATA MANAGER
DATA LIBRARIAN
DATA OFFICER
…

■ Bibliothèque
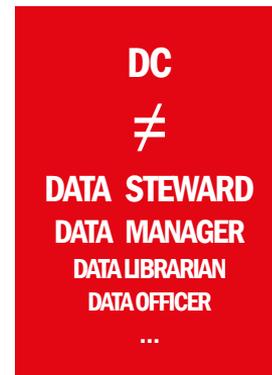
A DCs community
Pros & Cons

INSTITUTION

- Leverage on **motivated** persons
- Direct **influence** on practices
- Fastlane to **support** scholars

- Need (small) dedicated **HR**
- Difficult **measure** FAIR impact

DATA CHAMPIONS

- **Networking** to collaborations
- Improve **research**
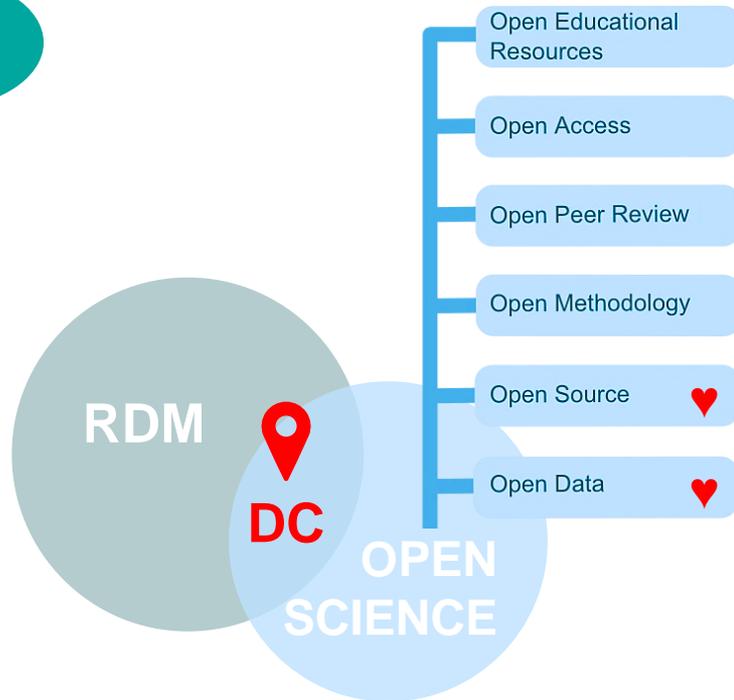- Peer **learning**
- **Discover** tools
- Boost **CV**s

- Need some (little) **availability**
- Need **visibility**

■ Bibliothèque

RDM → FAIR → Open Data!
Good practices already exist

AVERAGE MATURITY 78%

Policy • Portal • Impact • Quality • EU average

EU27 Open Data Maturity Report 2020

Open Educational Resources

Open Access

Open Peer Review

Open Methodology

Open Source ♥

Open Data ♥

RDM

DC

OPEN SCIENCE

Bibliothèque