

## Data Management in Research & Development, March 2, 2021

«Open Data and FAIR», Experience at BFH Institute for Medical Informatics

Serge Bignens

Institute for Medical Informatics I4MI BFH, Board Member MIDATA Cooperative

# Agenda

Open Data & FAIR Principles

Health Data

Data Reuse and Data Protection

MyData Principles

The Law

Our Experience for Solving the Challenges:

- ▶ MIDATA Data Governance and Infrastructure
- ▶ Example 1 *Ally Science* project
- ▶ Example 2 *Corona Science* project

# Open Data and FAIR Principles

## Open Data

Open data is publicly available data that can be universally and readily accessed, used, and redistributed free of charge. Open data is released in ways that protects private, personal or proprietary information. It is structured for usability and computability.

Source: Denecke, K. (2020). Biomedical Standards and Open Health Data. . In: D. Waltemath et al. (eds.) Reference Module in Biomedical Sciences. Elsevier, 2020, doi:10.1016/b978-0-12-801238-3.11527-2

# Open Data in Switzerland



## Opendata.ch

- ▶ is the swiss chapter from the [Open Knowledge Foundation](#), and is mainly active for e-government data



## SNSF:

- ▶ Research data should be freely accessible to everyone – for scientists as well as for the general public. Since October 2017, researchers have to include a data management plan (DMP) in their funding application for most of the funding schemes. At the same time, the SNSF expects that data generated by funded projects are publicly accessible in digital databases **provided there are no legal, ethical, copyright or other issues.**

## Open Data in Switzerland



**Swisstopo data are now available free of charge as Open Government Data in [geo.admin.ch](https://www.geo.admin.ch) geo- and download services.**

From **March 1, 2021**, the Federal Office of Topography swisstopo will make its official digital data and services available online free of charge as Open Government Data (OGD). As of now, no registration is required to download the data or use the geoservices WMS, WMTS and all services of the RESTful API.

Source: <https://www.geo.admin.ch/de/home.detail.news.html/geo-internet/news2021/news20210301.html>

*(Accessed on: March 1, 2021)*

## FAIR 4 Principles



**Findable**

**Accessible**

**Interoperable**

**Reusable**

Source: <https://www.go-fair.org/fair-principles/> (Accessed on: March 1, 2021)

# FAIR 4 Principles



## Findable

The first step in (re)using data is to find them. **Metadata and data should be easy to find for both humans and computers.** Machine-readable metadata are essential for automatic discovery of datasets and services, so this is an essential component of the [FAIRification process](#).

- ▶ [F1. \(Meta\)data are assigned a globally unique and persistent identifier](#)
- ▶ [F2. Data are described with rich metadata \(defined by R1 below\)](#)
- ▶ [F3. Metadata clearly and explicitly include the identifier of the data they describe](#)
- ▶ [F4. \(Meta\)data are registered or indexed in a searchable resource](#)

Source: <https://www.go-fair.org/fair-principles/> (Accessed on: March 1, 2021)



# FAIR 4 Principles



## Accessible

Once the user finds the required data, she/he needs to know **how can they be accessed**, possibly including **authentication and authorisation**.

- ▶ A1. (Meta)data are retrievable by their identifier using a standardised communications protocol
- ▶ A1.1 The protocol is open, free, and universally implementable
- ▶ A1.2 The protocol allows for an authentication and authorisation procedure, where necessary
- ▶ A2. Metadata are accessible, even when the data are no longer available

Source: <https://www.go-fair.org/fair-principles/> (Accessed on: March 1, 2021)

# FAIR 4 Principles



## Interoperable

The data usually need to be integrated with other data. In addition, the data need to **interoperate with applications** or workflows for analysis, storage, and processing.

- ▶ I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- ▶ I2. (Meta)data use vocabularies that follow FAIR principles
- ▶ I3. (Meta)data include qualified references to other (meta)data

Source: <https://www.go-fair.org/fair-principles/> (Accessed on: March 1, 2021)

# FAIR 4 Principles



## Reusable

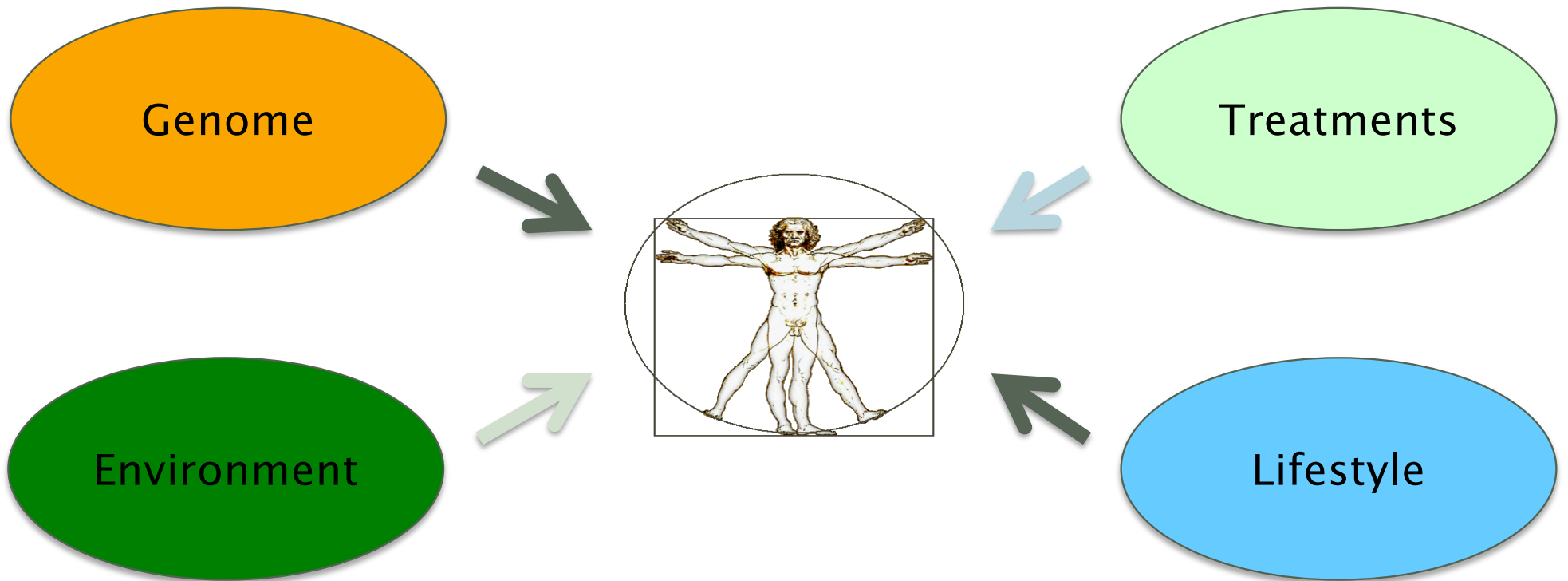
The ultimate goal of FAIR is to optimise the reuse of data. To achieve this, metadata and data should be well-described so that they **can be replicated and/or combined in different settings**.

- ▶ R1. (Meta)data are richly described with a plurality of accurate and relevant attributes
- ▶ R1.1. (Meta)data are released with a clear and accessible data usage license
- ▶ R1.2. (Meta)data are associated with detailed provenance
- ▶ R1.3. (Meta)data meet domain-relevant community standards

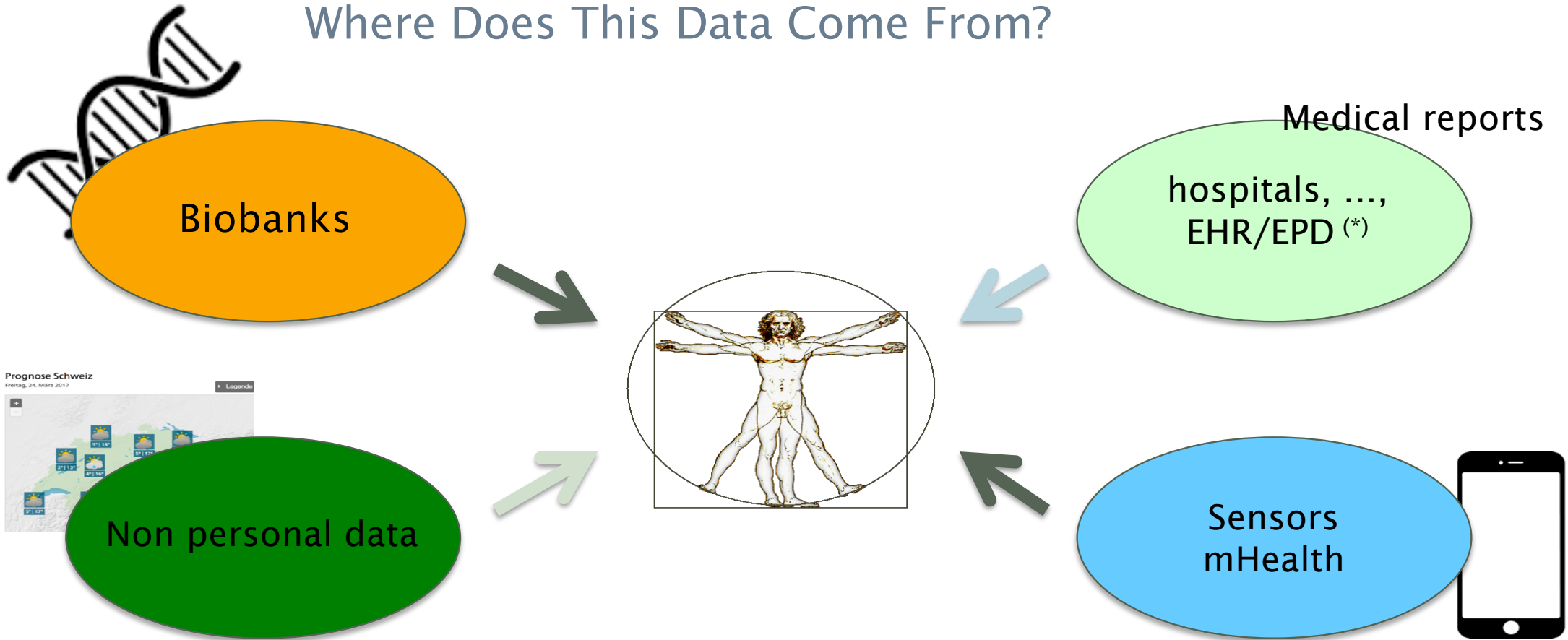
Source: <https://www.go-fair.org/fair-principles/> (Accessed on: March 1, 2021)

Data

## Determinants of Our Health



# Where Does This Data Come From?



(\*) EHR = electronic health record  
EPD = elektronisches Patientendossier

## Is Personal Data Equal to Personal Data ?

Well ... not exactly. There are:

- ▶ *“personal data»*  
e.g., Name, Surname
  
- ▶ *“sensitive personal data”*  
e.g., Diagnostic list, Medication list, Income, DNA, Religion

FADP Art 6 “Additionally, consent must be given expressly in the case of processing of sensitive personal data or personality profiles.”

# Two Dimensions



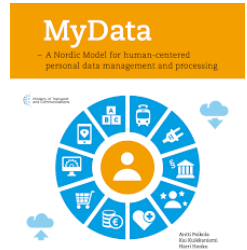
## Two Dimensions

Maximizing the (secondary) usage of data

and/or

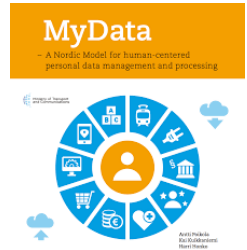
Maximizing the protection of the personal data

## MyData Targets Both



[MyData Global](#) is an international nonprofit association. MyData originated in Finland.

The purpose of MyData Global is to **empower individuals by improving their right to self-determination regarding their personal data**. It is based on the MyData [Declaration](#).



# MYDATA 6 Principles

## Human-centric control of personal data

- ▶ Individuals (...) understand and effectively control who has access to data about them and how it is used and shared. Individuals to be **empowered to give, deny or revoke their consent** to share data based on a clear understanding of why, how and for how long their data will be used.

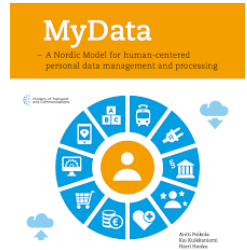
## Individual as the point of integration

- ▶ The value of personal data grows exponentially with their diversity; however, so does the threat to privacy. This contradiction can be solved if individuals become the “hubs” where, or through which **cross-referencing of personal data** happens.

## Individual empowerment

- ▶ We want individuals to be able to **securely manage their personal data** in their own preferred way. We believe that these are the preconditions for fair and beneficial data-based relationships.

Source: <https://mydata.org/declaration/> (Accessed on: March 1, 2021)



## MYDATA 6 Principles

### Portability: access and re-use

- ▶ The portability of personal data, that allows individuals to obtain and reuse their personal data for their own purposes and across different services, is the key to make the shift from data in closed silos to data which become reusable resources.

### Transparency and accountability

- ▶ Organizations that use a person's data should say what they do with them and why, and should do what they say.

### Interoperability

- ▶ Eliminating the possibilities of data lock-in. Open APIs, protocols, applications and infrastructure, so that all personal data are **portable and reusable**, without losing user control.

Source: <https://mydata.org/declaration/> (Accessed on: March 1, 2021)

# The Law (in Switzerland)

# Most Relevant Swiss Laws

## Federal Act on Data Protection

(FADP)

of 19 June 1992 (Status as of 1 March 2019)

**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)**

Sources:

[https://www.fedlex.admin.ch/eli/cc/1993/1945\\_1945\\_1945/en](https://www.fedlex.admin.ch/eli/cc/1993/1945_1945_1945/en)

<https://www.fedlex.admin.ch/eli/cc/2013/617/en> (the current valid version 1.2.2021 is not yet published in English)

## Federal Act on Research involving Human Beings

(Human Research Act, HRA)

of 30 September 2011 (Status as of 1 January 2020)

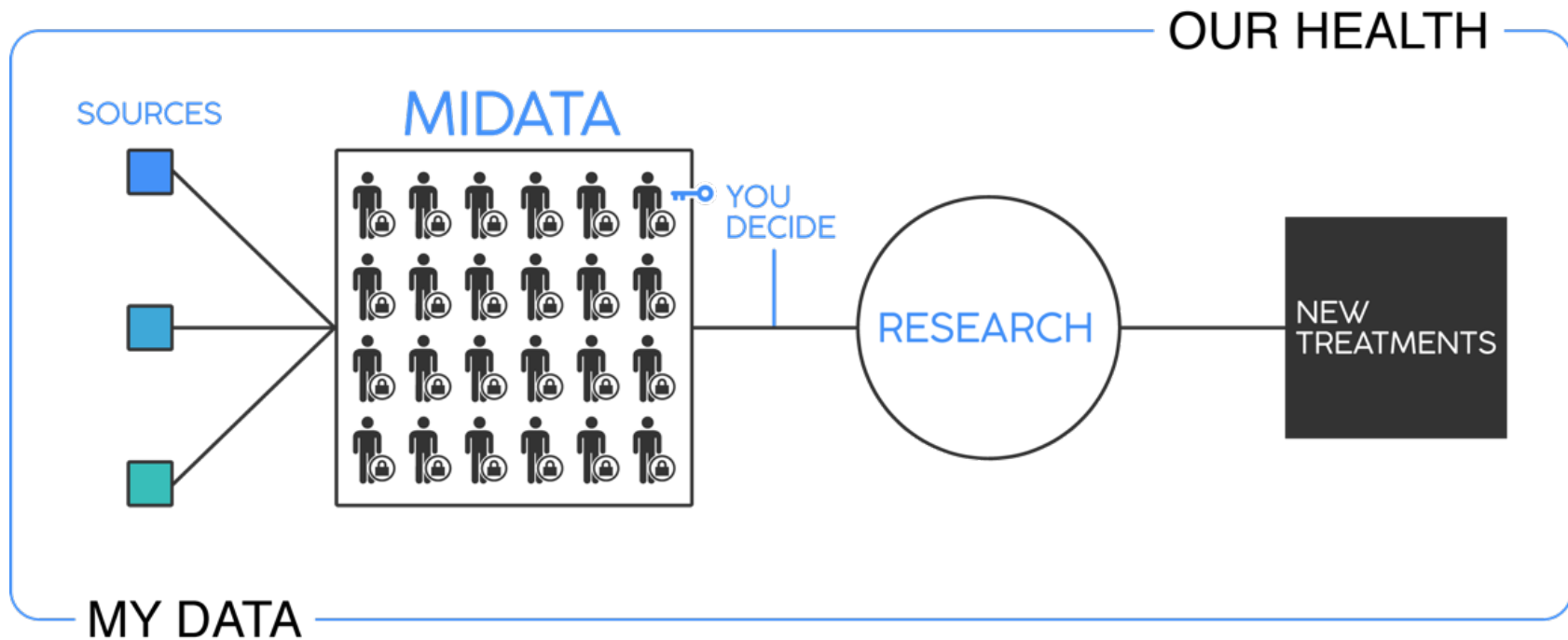
HRA Art. 3

- h. *Coded biological material and coded health-related personal data* means biological material and data linked to a specific person via a code;
- i. *Anonymised biological material and anonymised health-related data* means biological material and health-related data which cannot (without disproportionate effort) be traced to a specific person;

# Our Experience for Solving the Challenges

# MIDATA «my data, our health»

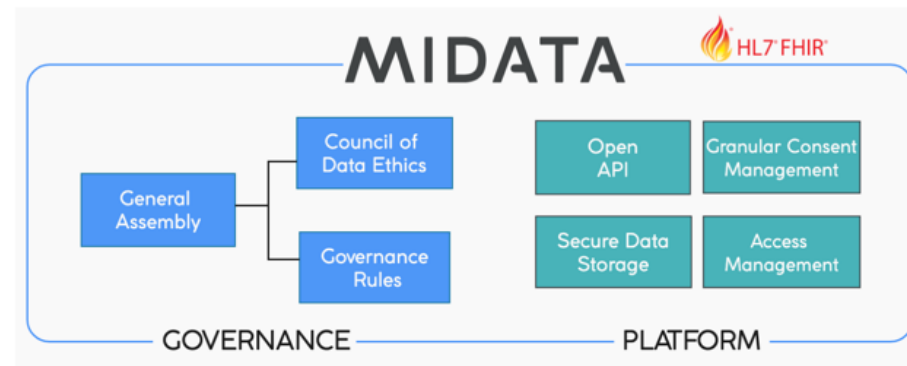
## Citizen-Centered Data Aggregation and Sharing



[www.midata.ch](http://www.midata.ch)



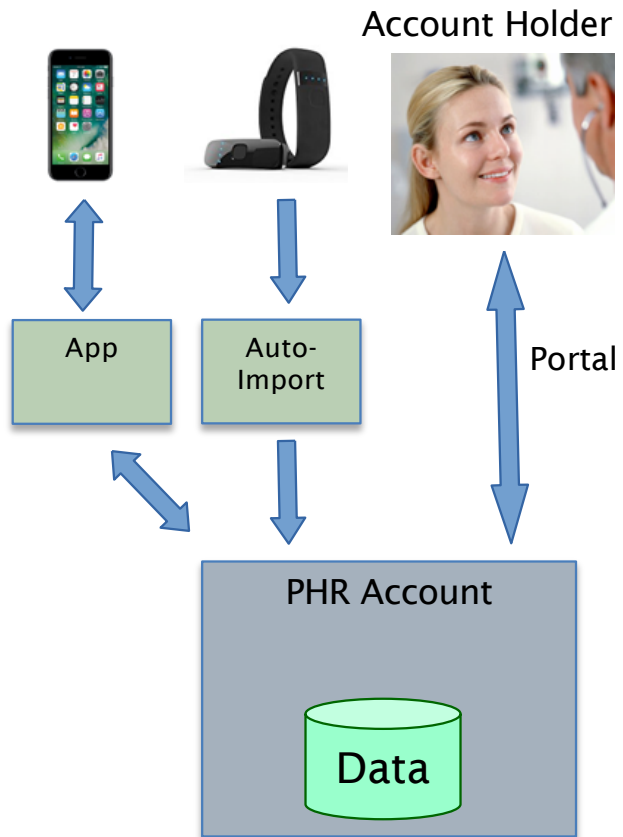
# Data Governance of the MIDATA Personal Health Data Storage Platform



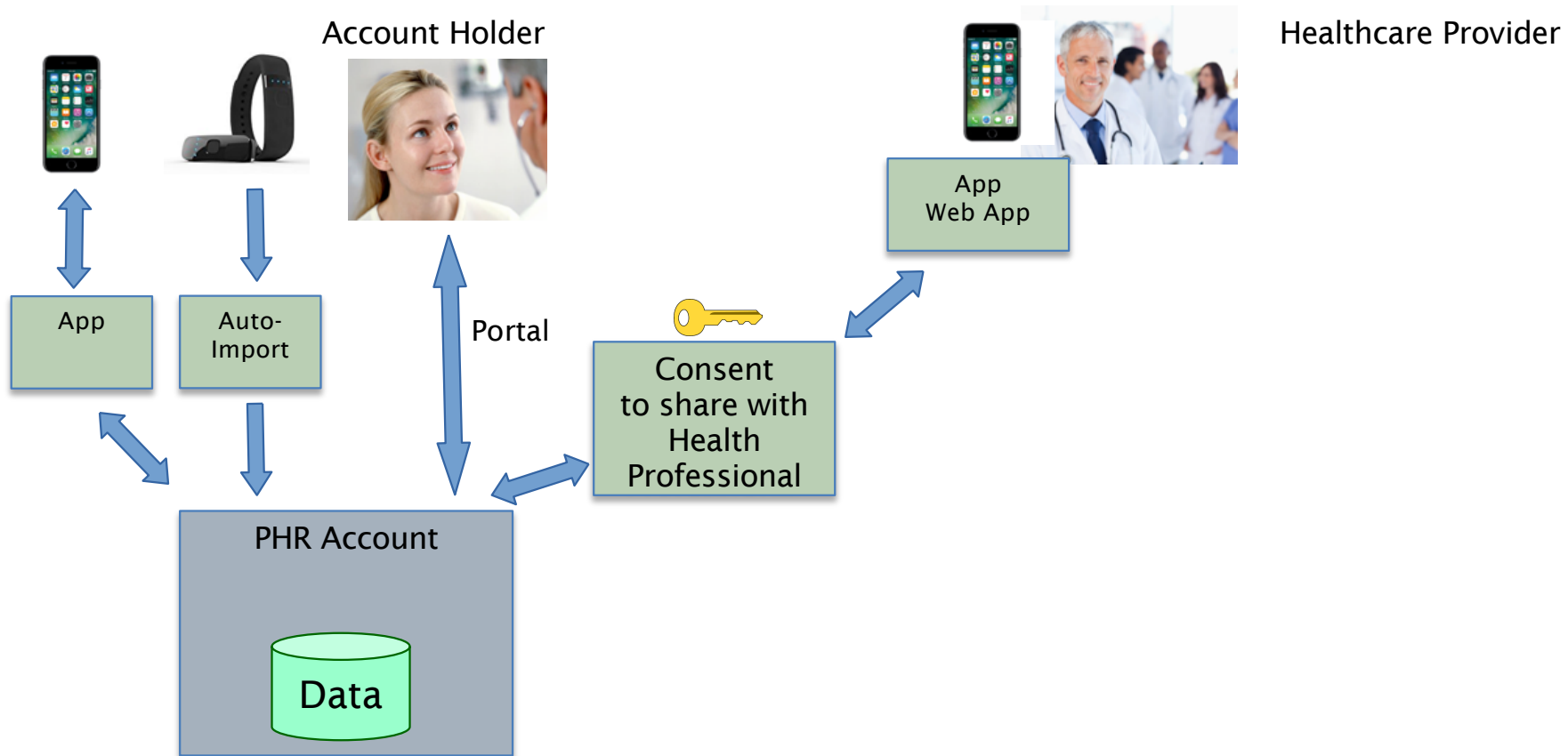
## Elements of governance

- ▶ Organizational form: non-profit cooperative
  - ▶ Cooperative members govern cooperative
  - ▶ Data account holders are asked to become members
- ▶ Organizational constitution: articles of association
  - ▶ Organs of governance
    - ▶ General assembly
    - ▶ Council of data ethics
  - ▶ Data account holders' rights are codified
    - ▶ In the articles of association
    - ▶ In the general terms and conditions
    - ▶ In the privacy policy

# MIDATA Platform, Data Flow and Data Management

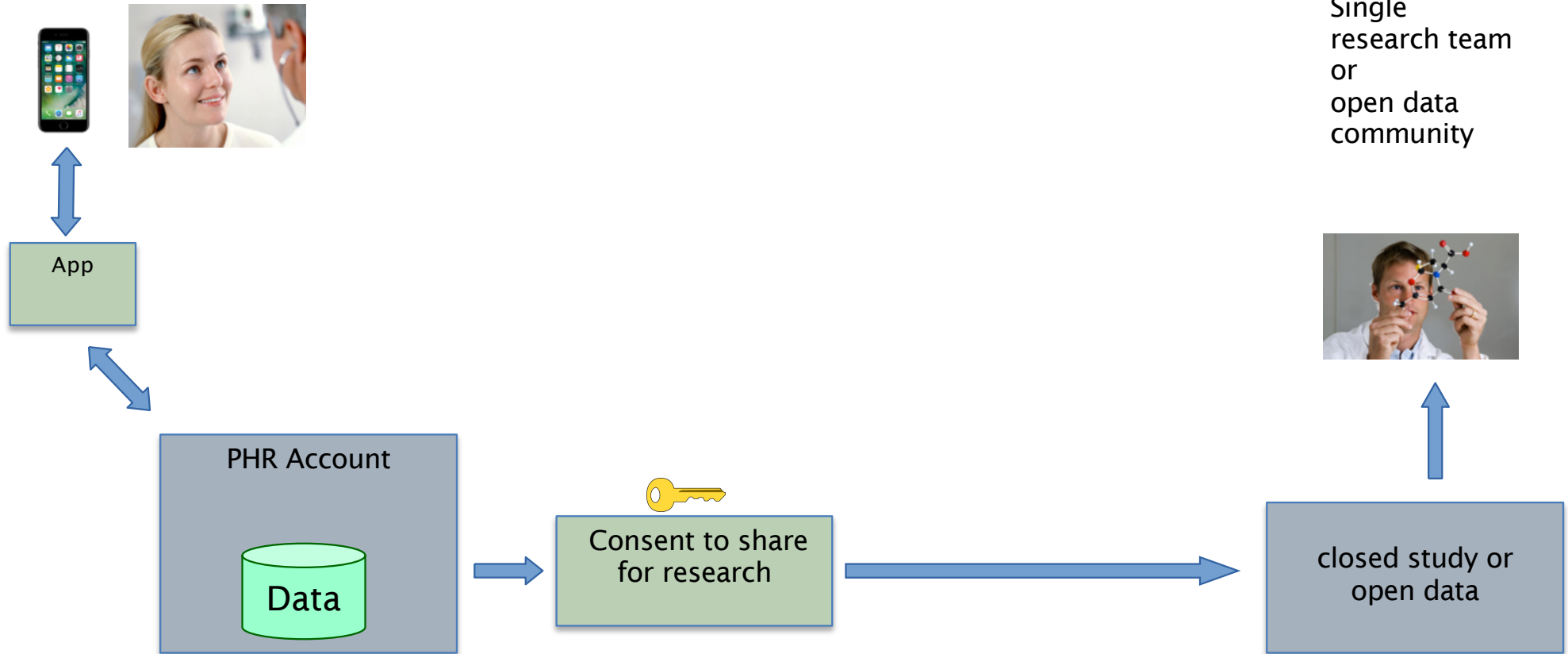


# MIDATA Platform, Data Flow and Data Management



# MIDATA Platform, Data Flow and Data Management

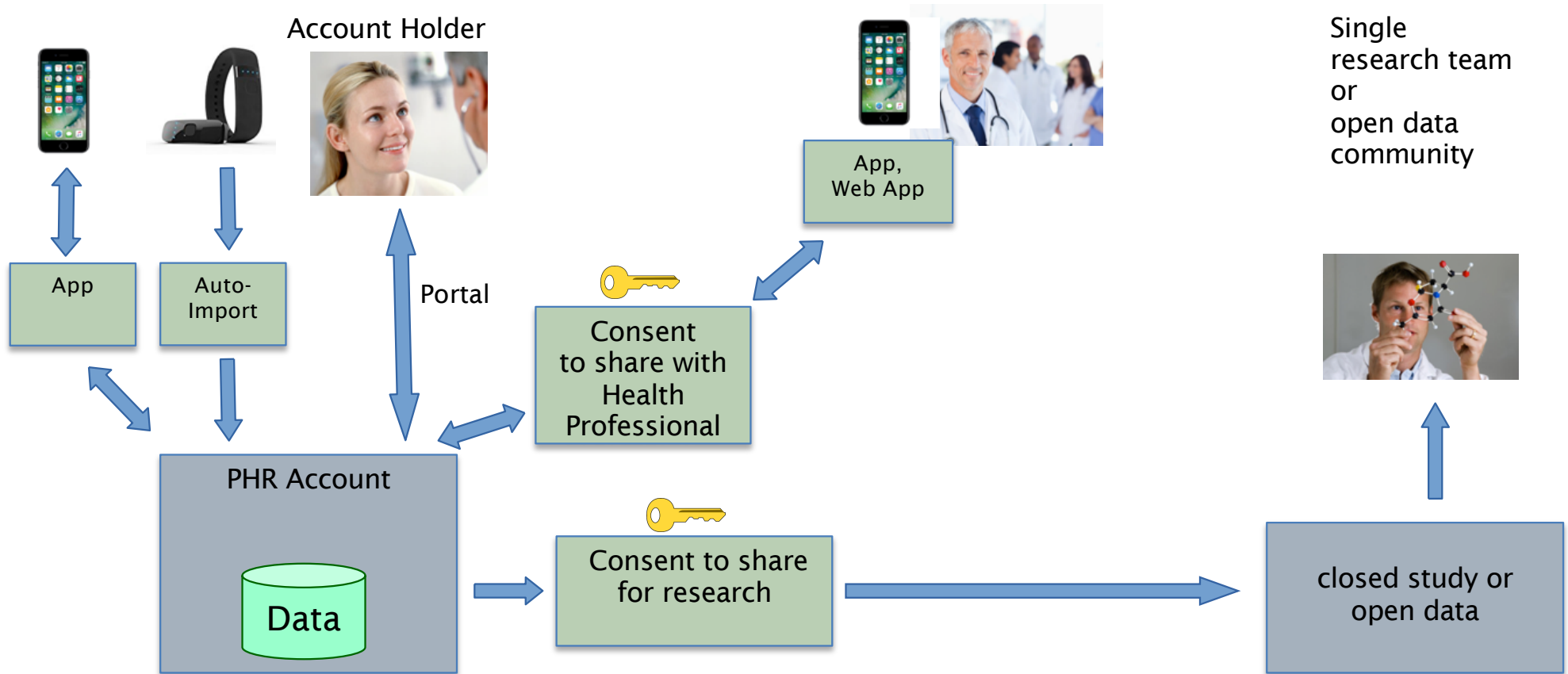
Account Holder



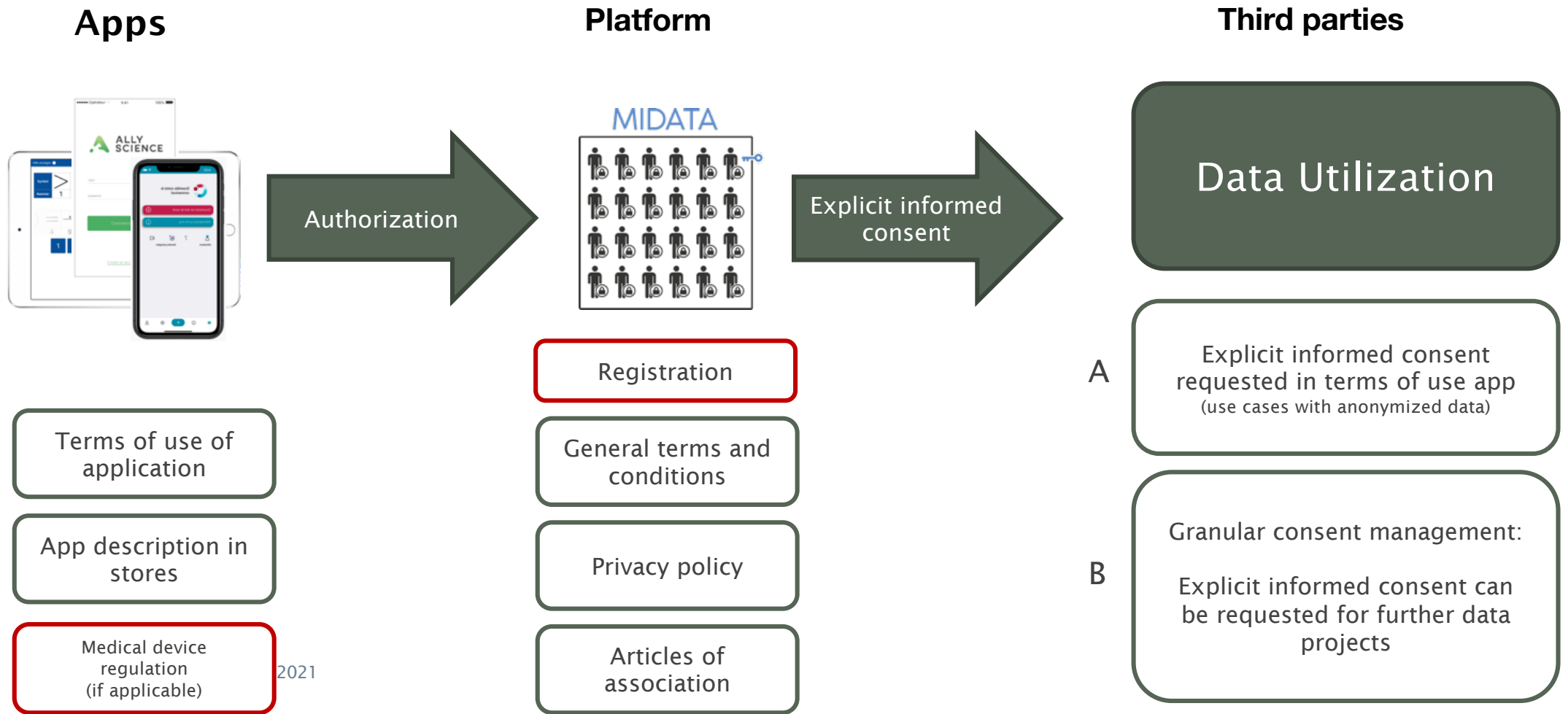
# MIDATA Platform, Data Flow and Data Management

Healthcare Provider

Account Holder



# Data Management and Access Control



# App-Associated Use Cases

## Example: Corona Science

Terms of use of application

<https://ch.midata.coop/#/portal/terms/corona-science-terms-of-use--1.0>

App description in stores

<https://apps.apple.com/ch/app/corona-science/id1504035002>

<https://play.google.com/store/apps/details?id=com.coronascience>

Registration

<https://ch.midata.coop/#/portal/login>

General terms and conditions

<https://ch.midata.coop/#/portal/terms/midata-terms-of-use>

Privacy policy

<https://ch.midata.coop/#/portal/terms/midata-privacy-policy>

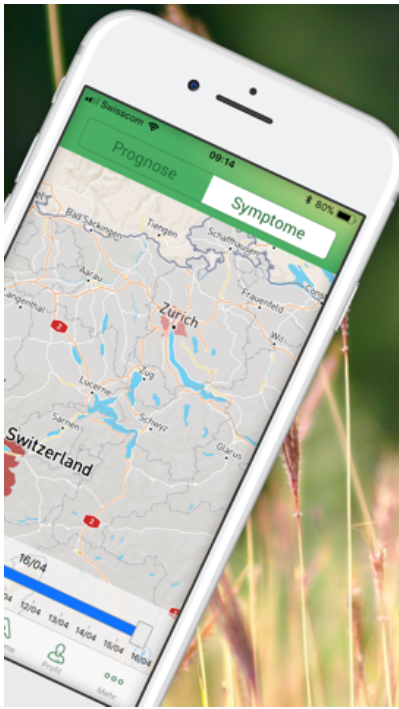
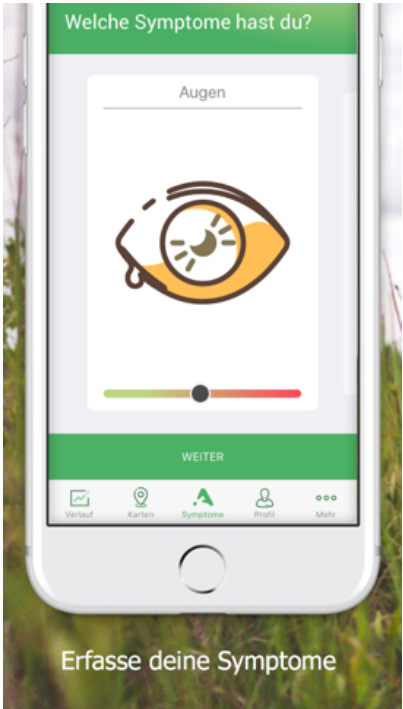
Articles of association

[https://www.midata.coop/wp-content/uploads/2019/08/MIDATA\\_Statuten\\_20190626\\_EN.pdf](https://www.midata.coop/wp-content/uploads/2019/08/MIDATA_Statuten_20190626_EN.pdf)

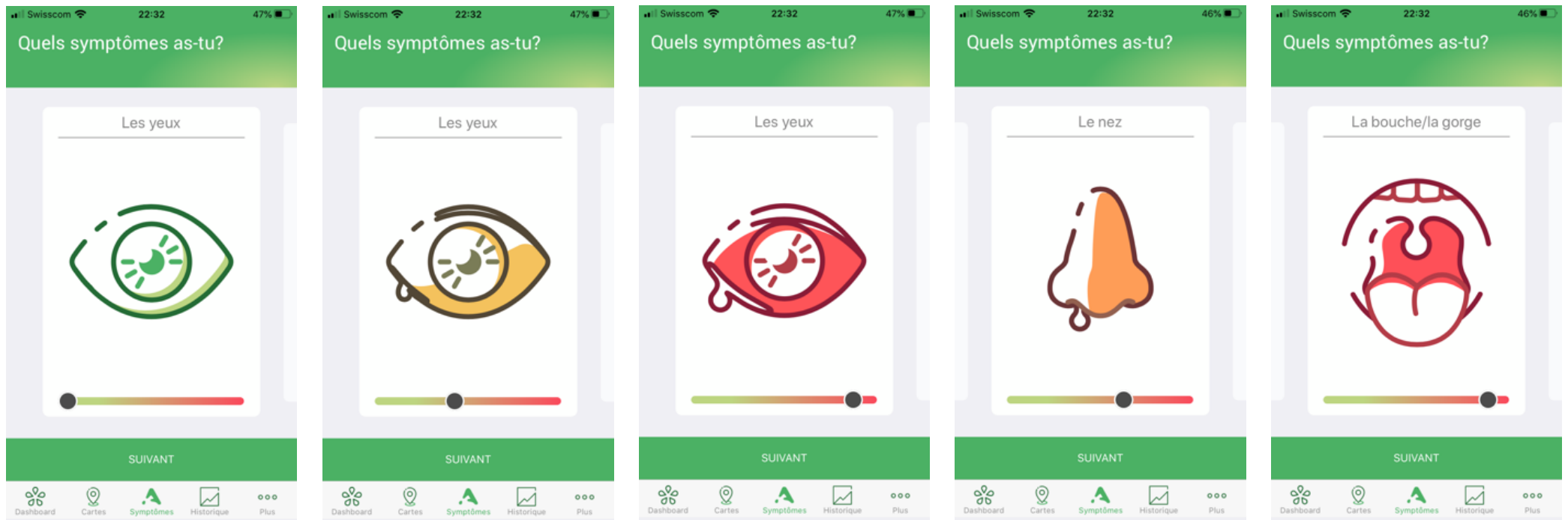
# Our First Large Data Project; 10'000 Participants



Ermögliche die grösste  
Allergie-Studie  
der Schweiz



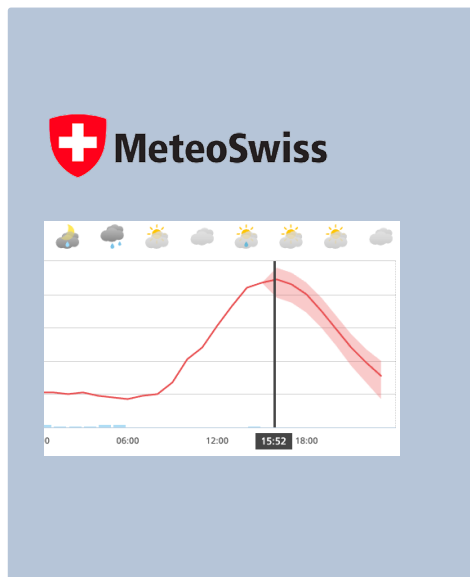




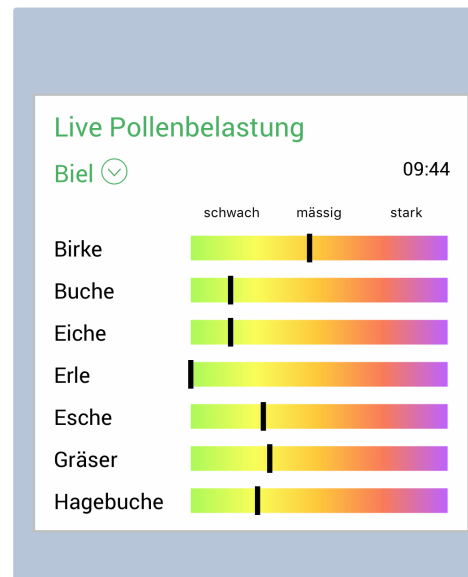
[www.allyscience.ch](http://www.allyscience.ch)

# Aggregating Environment, Symptoms and Treatment Data

Pollen forecast



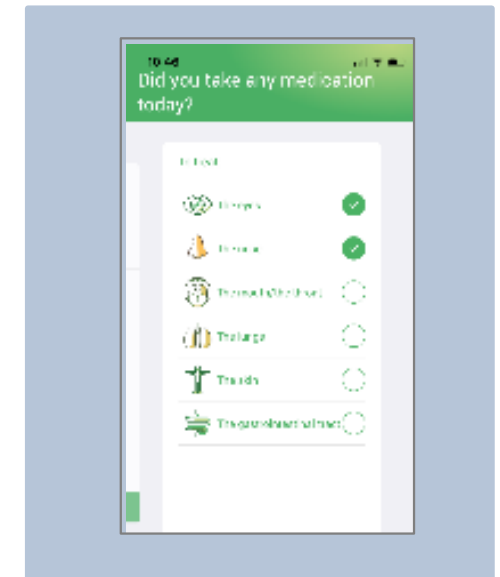
Pollen real time measures



Patient reported outcomes



Medication



# Ally Science Data Sharing with MIDATA

Account Holder



Healthcare Provider



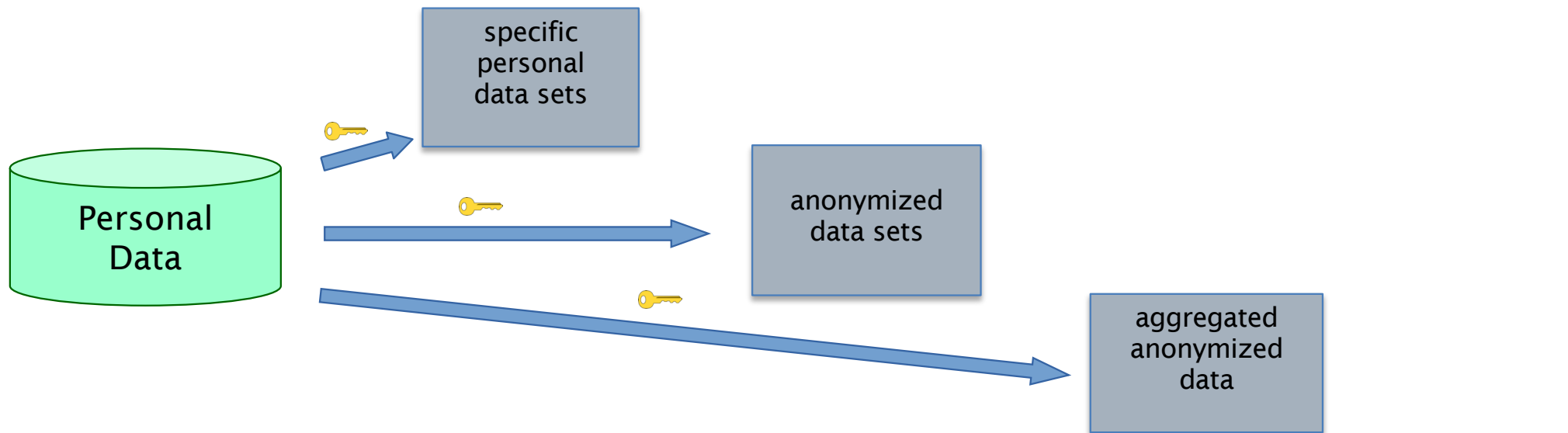
USZ Research team



Ally Science user community



Open data



## Ally Science: Lessons learned

>80% of people having downloaded the app, used the app and accepted to open an MIDATA account and accepted the explicit dynamic consent to share data with the ally science community and with the researchers of USZ. Which shows a very high acceptance rate for those 4 steps.

5000 participants registered within the first week

Whole project from study design, financing, specification, implementation, guide lines, tests, security audit, translation in 6 languages, communication concept, video support and launch within 4 months.

This:

## Storage

### HRO Art. 5, Para. 2, Let. b

u<sup>b</sup>

b  
UNIVERSITÄT  
BERN

*b. prevent unauthorised or accidental disclosure, alteration, deletion and copying of the health-related personal data*

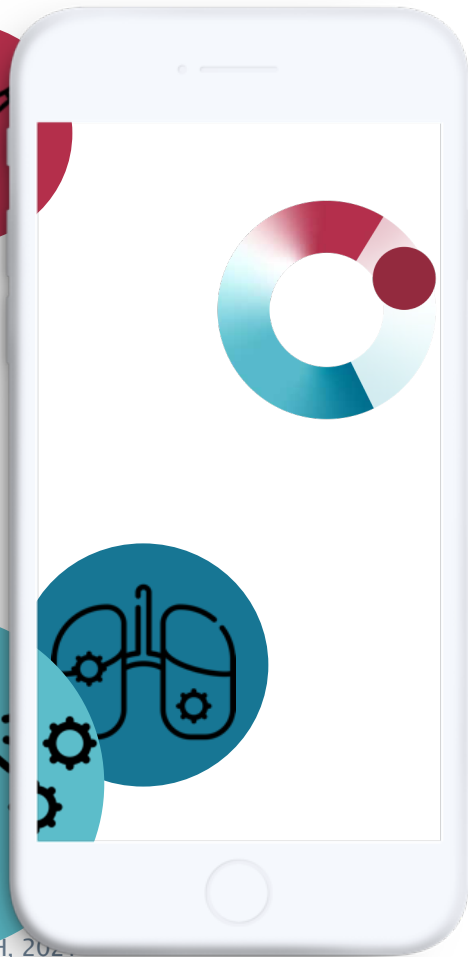
- Location of storage → central storage/server (own computer???)
- Access → user management
- Internal standards and defined processes → quality management (system) and standard operation procedures
  
- Prospective studies (data entry): dedicated study management system with a data model
  - Frontend i.e. electronic case report forms
  - Backend i.e. relational database

Source: presentation Sven Trelle March 2<sup>nd</sup> 2021

Is solved by configuring MIDATA + BFH/I4MI data management framework

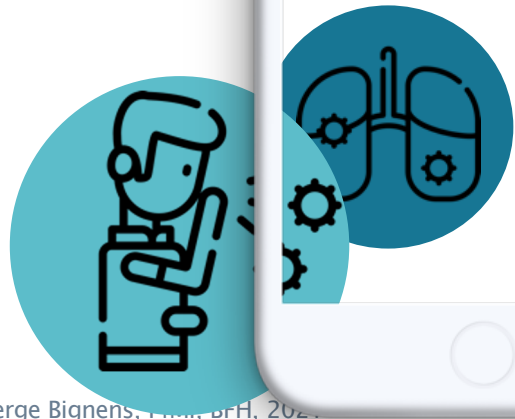
# Our Second Large Data Project and First as Open Data

Join this ongoing project!



## CORONA SCIENCE

Citizen-reported data for the Open Data community

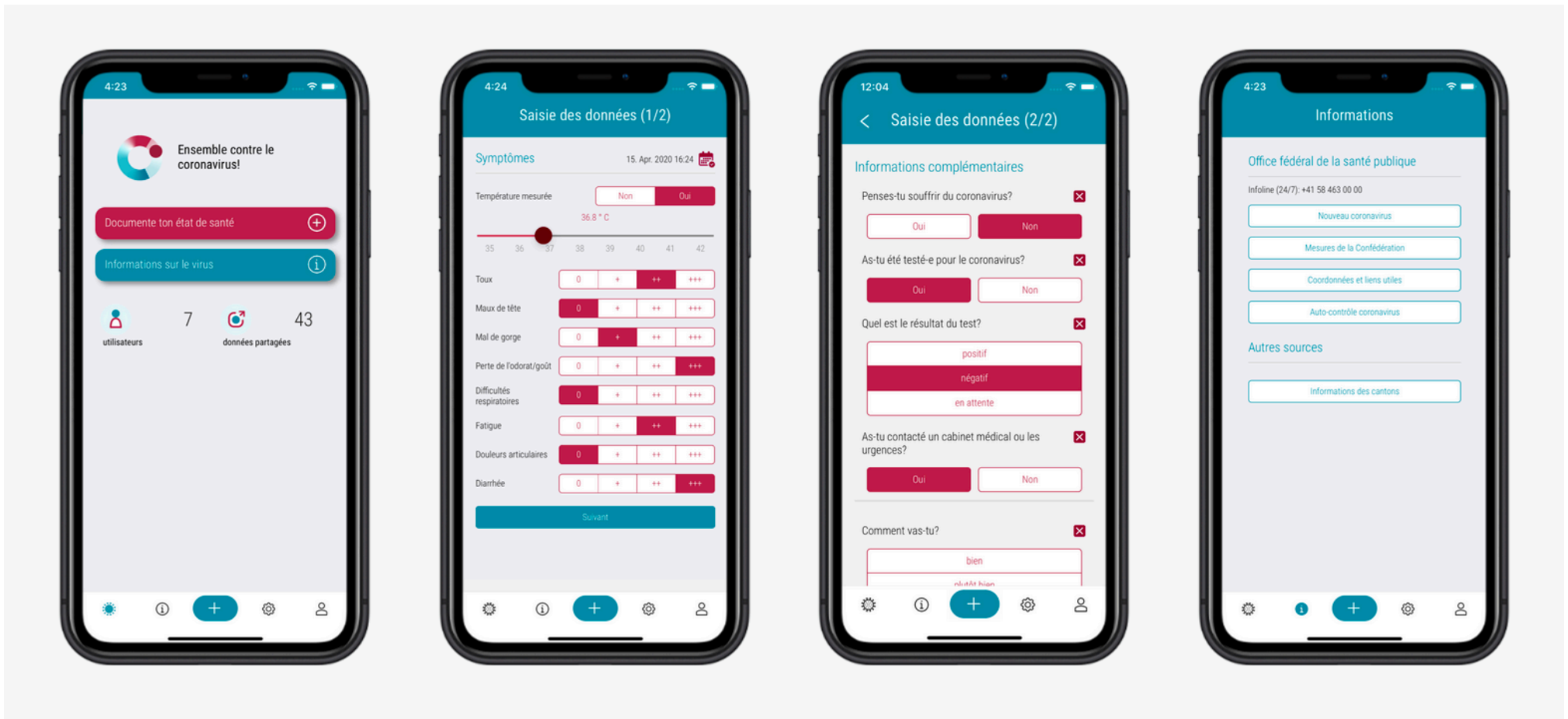


Serge Bignens, FH, 2021



[www.coronascience.ch](http://www.coronascience.ch)



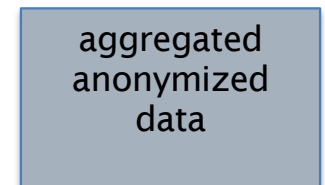
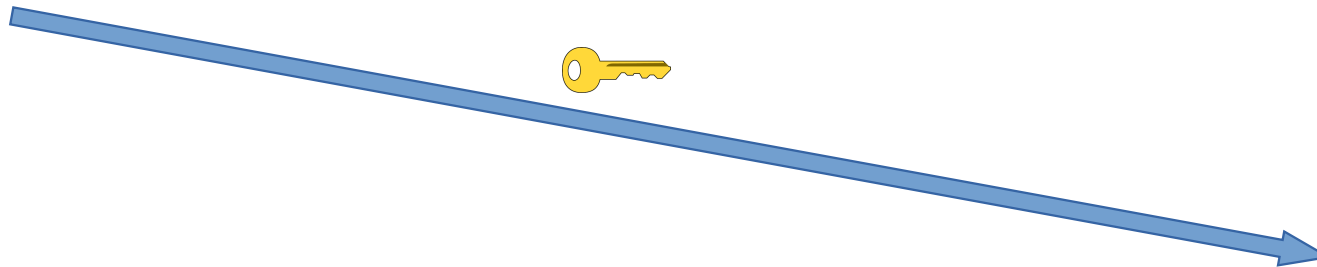
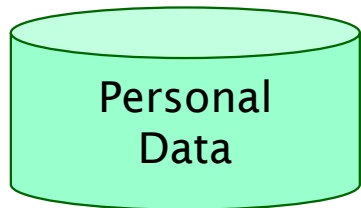


# Corona Science Data Sharing with MIDATA

Account Holder



Open data





## Corona Science: Lessons learned

10'000 participants registered within the first week = people trust our framework.

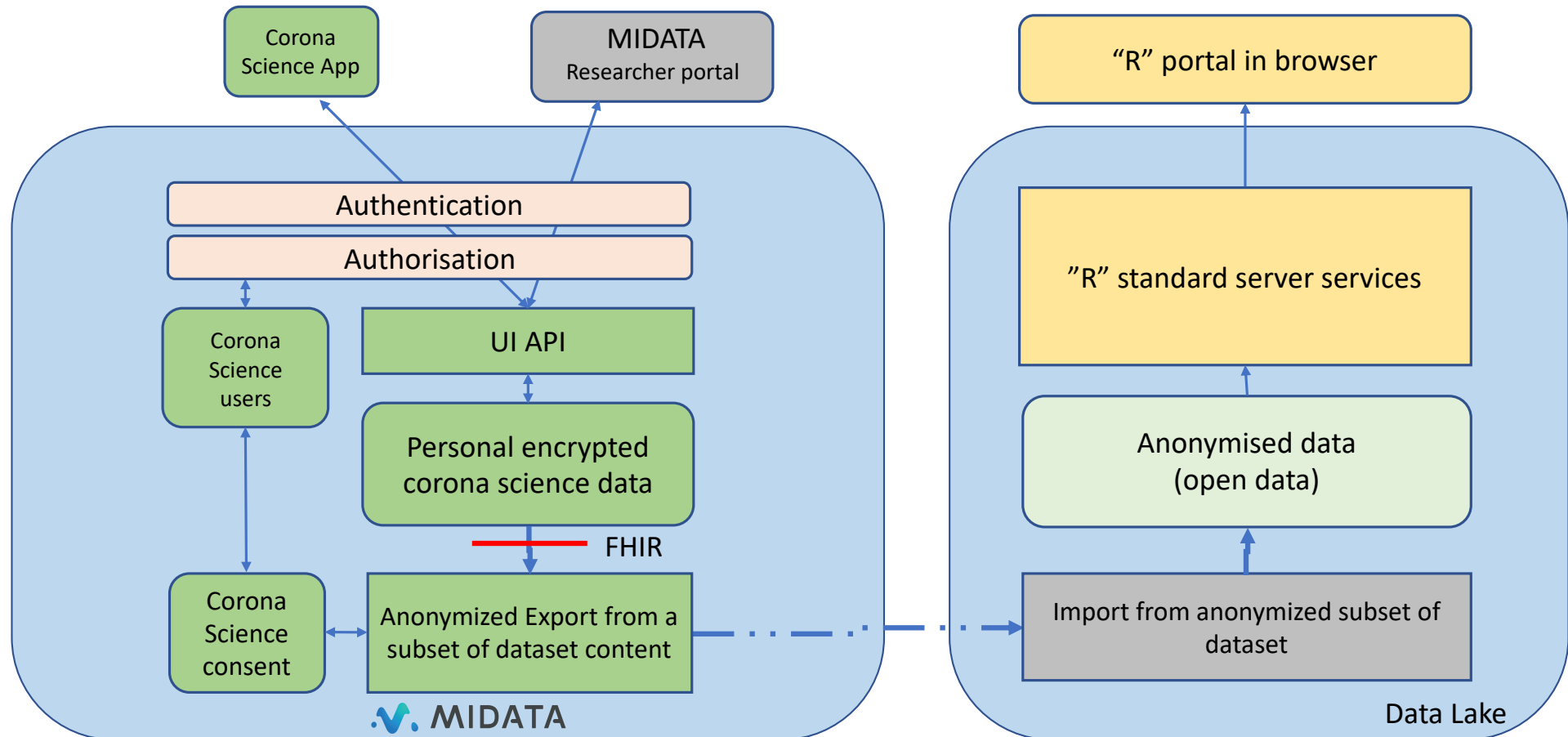
First corona app on apple store and google play store in Mid April 2020.

Not so easy to find repository for open data, as it is still a relatively new concept when it goes to real implementation.

Whole project, definition, setup, development, deployment in 3.5 weeks, and this thanks reuse of building blocks:

- ▶ MIDATA data governance framework
- ▶ Available translation of documentation framework
- ▶ MIDATA IT platform, yearly audited for security
- ▶ Mobile App Framework from BFH I4MI
- ▶ Cooles team ;-)
- ▶ Access to external expertise (communication, social network, legal)

## And for the techies among you ;-)



Thanks for your attention

Time to enter the interaction phase of the workshop!



[serge.bignens@bfh.ch](mailto:serge.bignens@bfh.ch)



[@Sbignens](https://twitter.com/Sbignens)



[@sergebi](https://www.instagram.com/sergebi)



<https://www.linkedin.com/in/sergebignens/>

[www.midata.coop](http://www.midata.coop)

[www.allyscience.ch](http://www.allyscience.ch)

[www.coronascience.ch](http://www.coronascience.ch)