HEALTH DATA HUB

Presentation of the French Health Data Hub & outlook for DIPoH
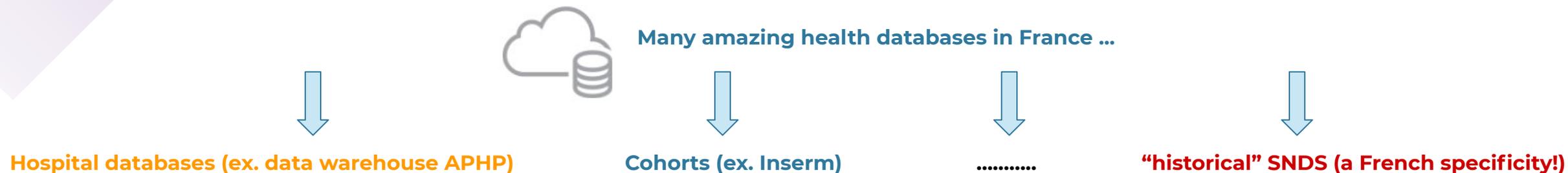June 3rd, 2021
Emmanuel Bacry
Health Data Hub / CNRS

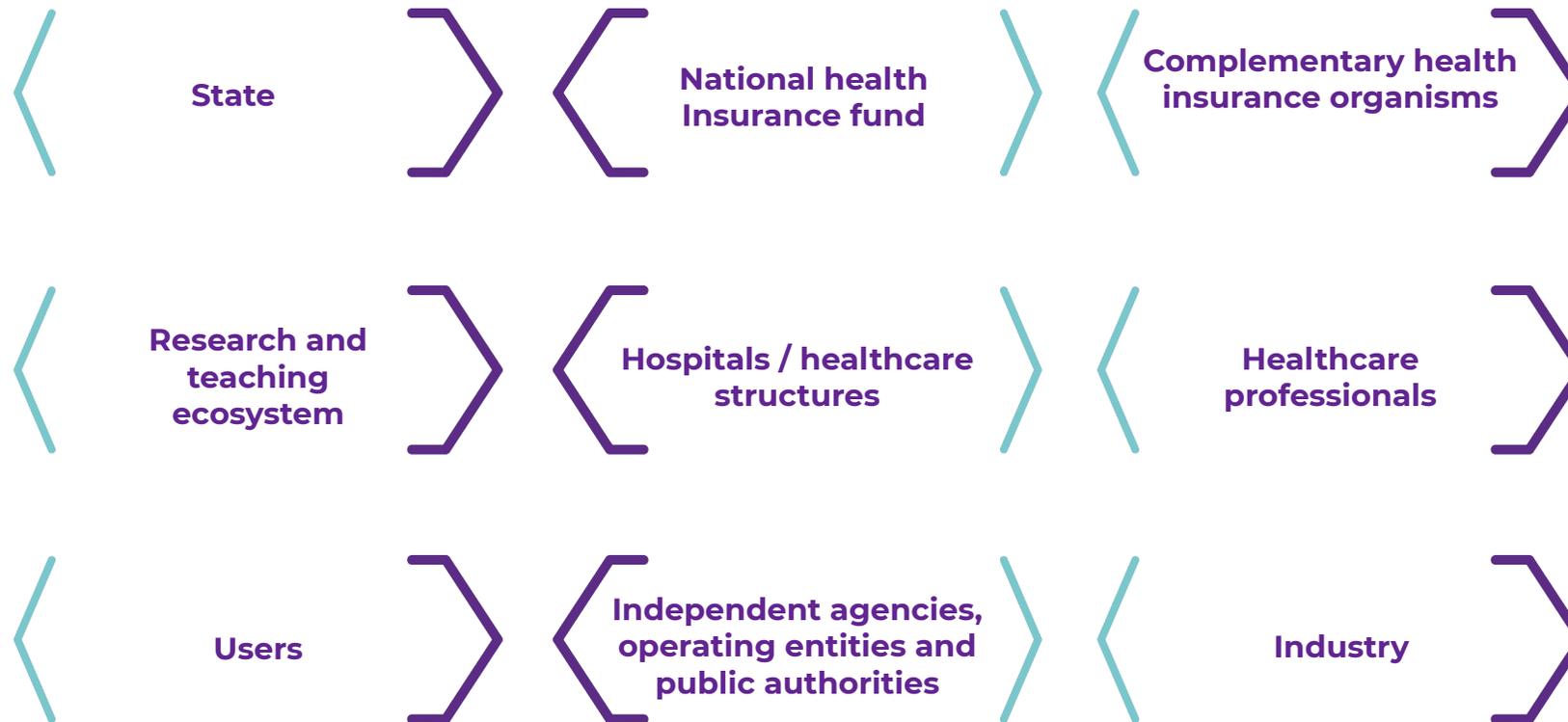# The Health Data Hub

Health databases in France

Many amazing health databases in France ...

Hospital databases (ex. data warehouse APHP)   Cohorts (ex. Inserm)   ....   "historical" SNDS (a French specificity!)

BUT Though many important research projects are already operating on them, they are still largely underexploited

Mainly because :

1. The data are scattered among multiple databases, little is known and not easily understood by "outsiders"

2. The gouvernance and procedures for accessing health data are complex and heterogeneous, especially because of the sensitivity of these data

3. There is lacking interoperability of health data

The Health Data Hub is …

… a **public interest group** created late 2019 and set out by ministerial decree, bringing together **56 stakeholders divided into 9 groups** :

State

National health Insurance fund

Complementary health insurance organisms

Research and teaching ecosystem

Hospitals / healthcare structures

Healthcare professionals

Users

Independent agencies, operating entities and public authorities

Industry

These entities represent the diversity of the health data ecosystem. Their representatives are part of the executive board. The HDH is mainly publicly funded.

The Health Data Hub aims to facilitate access to health data for public interest research.

Created late 2019, the Health Data Hub guarantees easy, unified, transparent, and secure access to health data to improve the quality of care and patient support.
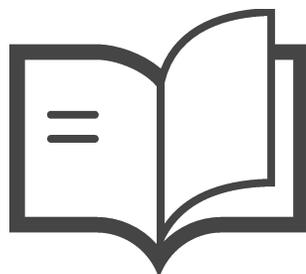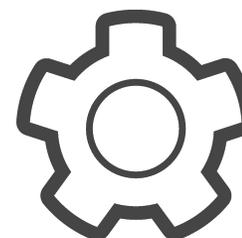Its mission is divided into four main pillars :

A **unique gateway**

A **state-of-the-art** and **secure platform**

A **documented data catalogue** built in a progressive manner

A **range of tools** to bring together key stakeholders in the sector

A data catalogue built in a progressive manner available for public interest research

Illustration - examples of databases that will be included in the Health Data Hub catalogue in 2021

"Historical" SNDS	Data from health insurance claims forms, hospital discharge and billing and medical causes of death
OSCOUR® database	The only French syndromic surveillance system based on emergency departments (Santé Publique France)
French Bone Marrow Transplant register	Database allowing the study of healthcare pathways between the announcement of the donation and the donation of bone marrow (Agence de Biomédecine)
Base of 500,000 angioplasties	Database from different centres to evaluate the impact of stents in real life (French Society of Cardiology)
I-share cohort	Database allowing to conduct studies on data related to a prospective cohort of 30K college students in France (University of Bordeaux)
and many more expected...	....

PURPOSE

The SNDS is a database that collects and makes available, for defined purposes, pseudonymized individual health data from the main medico-administrative sources.

KEY FIGURES

One of the biggest claims databases in the world
- Claims data of 65 million people
- 1.2 billion care sheets per year
- 11 million hospital stays per year = 250 Tb of structured data

Available on the Scala/Spark bigdata framework SCALPEL3 (//arxiv.org/pdf/1910.07045.pdf)

HEALTH DATA HUB

The Health Data Hub in 2020

PROJECT SUPPORT

1 Call for projects co-organized with the grand challenge of « Improving medical diagnosis with AI» in 2020

27 pilot projects   and   **9** Covid projects supported

+ with **20** partnerships under discussion with data controllers

SETTING UP THE ORGANISATION

55 collaborators

32 women, 23 men

36 recruitments in 2020

1 new working space

PROVISION OF THE TECHNOLOGICAL PLATFORM

2 versions of the platform in the production

3 projects on the platform

2 databases on the platform

1 ANSSI audit to certify high level of security (no critical vulnerabilities identified)

FEDERATING THE NATIONAL AND INTERNATIONAL ECOSYSTEM

**526** participants in the "Health data and collective intelligence" conference

547 data challenge participants, co-organised with French Society for Pathology (SPF)

438 participants in the « AI4Health » Winter School

9% of the EU budget was mobilised for France in the framework of the European Joint Action "Towards a European Health Data Space"

167 researchers in the affiliated researchers network

1250 members in the HDH meetup group
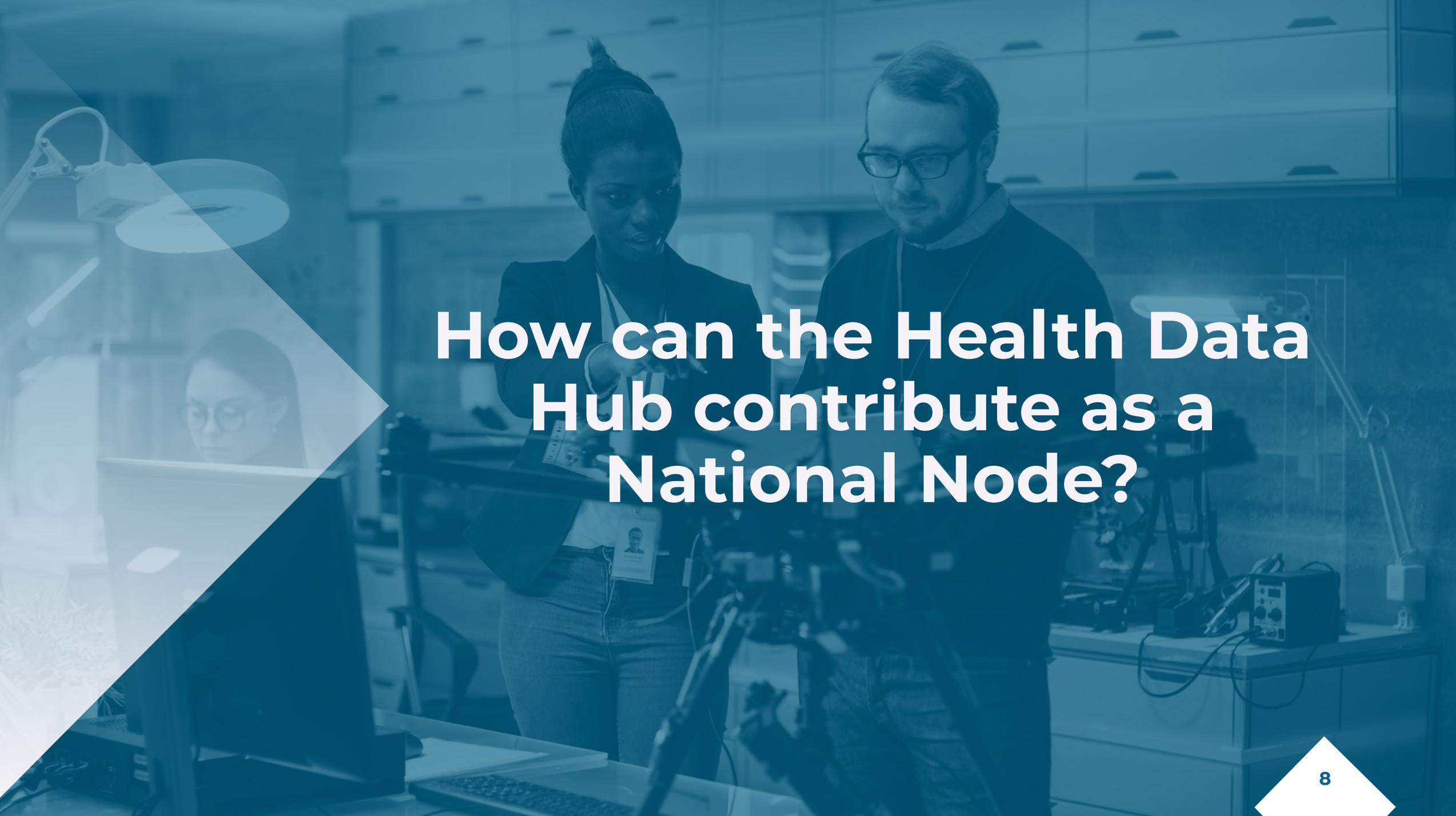
How can the Health Data Hub contribute as a National Node?

The Health Data Hub fulfills key success factors of a National Node

Visibility of data sources

Clear and unified gouvernance

Interoperability of data sources

→ Data pairing of different sources increases potential of data exploitation and increasing the depth and richness of information

→ Capability to exploit multiple resources also enables interdisciplinary research

→ Commitment to convert SNDS into OMOP-format in collaboration with EHDEN

Promote high security standards at scale

SUCCESS FACTORS of a national node

Strong potential to federate

Create a community of researchers and a health data ecosystem

Established scientific and ethics evaluation

Promoting high-quality research while respecting strict ethics standards

National nodes need to be part of a wider network of interconnected national nodes across the EU.

HEALTH DATA HUB

9

Why did the Health Data Hub join DIPoH?
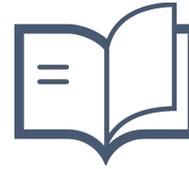
DIPoH: expectations and added value

Share its expertise & promote partnerships

Within the framework of DIPoH, the HDH could contribute by:

- sharing its **expertise/experience of having set up a single health data access point in France**

- **promoting and establishing bilateral and multilateral partnerships** for the exchange of good practices at the European level

Making data available throughout the EU

- The HDH has included information on its data sources in **the Health Information Portal**.

- The HDH will **make its data descriptions available** and contribute to the establishment of a **EU metadata catalogue.**

- The HDH will **publish via DIPoH English guidelines for accessing French data,** thus achieving its objective of allowing researchers from all over Europe to access French health data.

Setting up of multilateral conferences for international researchers

- The HDH will also be able to actively participate in the **setting up of conferences for European and international researchers and/or thematic meetings** (focus on AI, specific pathologies and other), **with the aim of promoting cross-border research collaborations.**
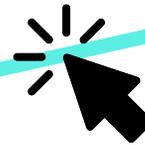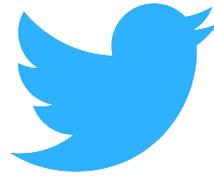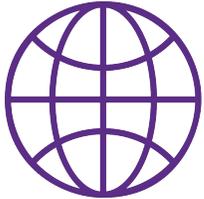
Good news! These guidelines have just been published on the HDH website.

Health Data Hub

Thanks

Contacts

Link to our website

Our engagement with civil society

Candidate

Our news