




HEALTH
DATA HUB

A woman in a dark blazer and jeans is pointing at a laptop screen in a laboratory setting. A man with glasses and a beard is standing next to her, looking at the screen. In the background, another woman is working at a computer. The scene is set in a modern laboratory with glass partitions and various equipment.

Roadmap for 2021



Health Data Hub: review and outlook

A word from the President and Vice-President

*"The crisis has illustrated **the central role played by data today**, and its essential contribution to improving our understanding and ability to anticipate, from a scientific perspective but also for the management of public policies. This situation has simply confirmed **the value and urgency of the Health Data Hub's mission**. In an extremely competitive international environment, it must maintain a high level of activity and meet the deadlines set for it."*



Bruno Maquart,
President of the HDH

*"Donating blood is all about helping others; so is sharing your health data. **Sharing it with your healthcare team enables you to receive better healthcare; sharing it with everyone improves the health care of all our fellow citizens.**"*



Gérard Raymond,
*President of France Assos Santé,
Vice-President of the HDH*

Editorial

One year after its creation, the Health Data Hub has achieved several milestones, in line with the commitments made in the Strategic Roadmap for 2020-2022.

2020 was marked by the **product launch**, in April, of the technological platform designed to provide authorised users with the data required for their projects. The first databases are already available, and the first users are now working with the data as part of the efforts to overcome the COVID-19 pandemic.

In addition, the Health Data Hub is **supporting data access for around forty projects**, twenty of which were selected following calls for projects. Several **events** were organised, including a medical imaging data challenge in partnership with the French Society of Pathology and a conference illustrating the benefits of using health data.

The Health Data Hub has also been chosen to coordinate the French response to the European Commission's joint action to create a "European Health Data Space". Finally, **work on developing the Health Data Hub's information policy for citizens was carried out with patient and healthcare system user associations.**

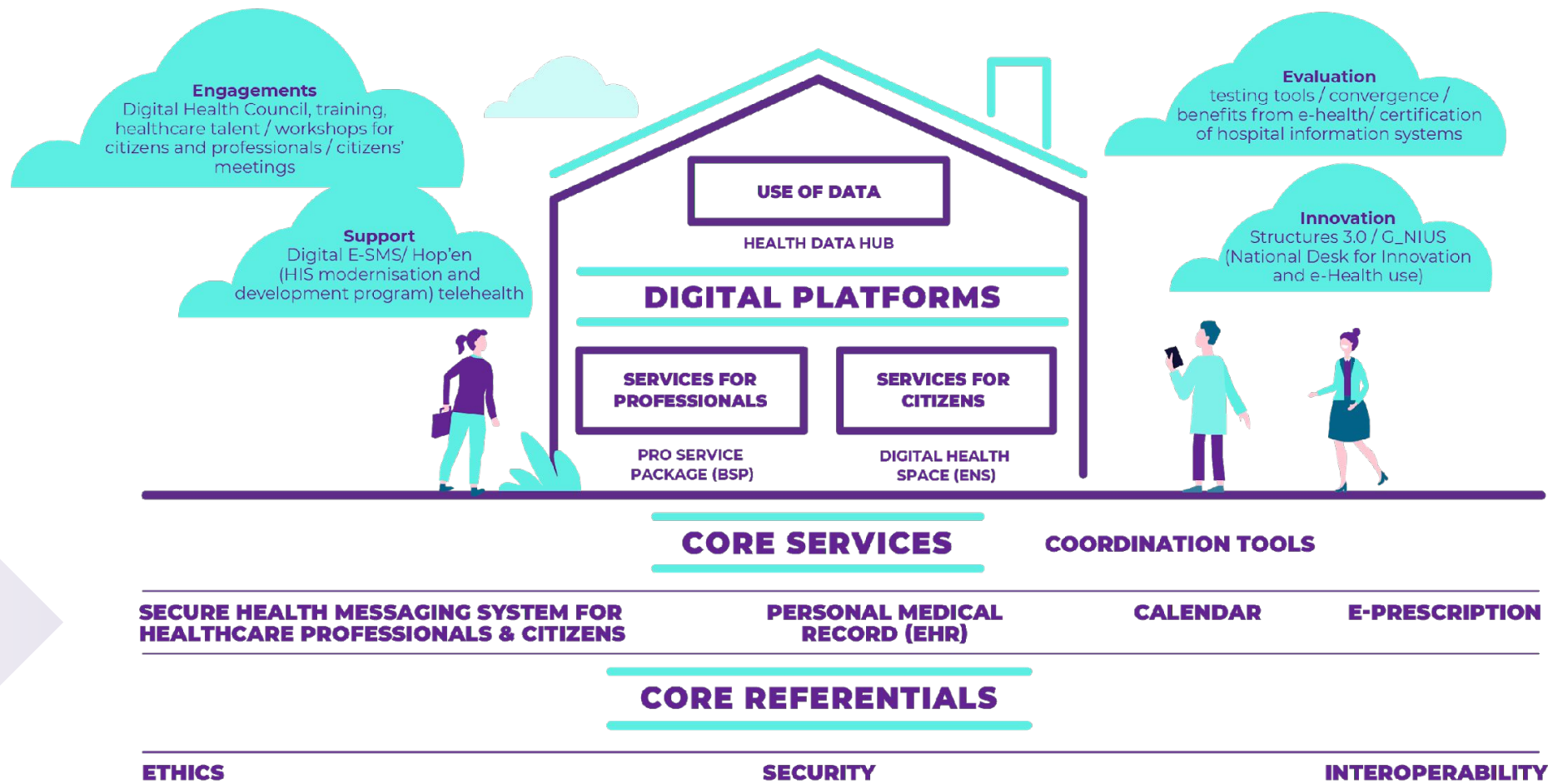
In 2021, the Health Data Hub will continue to roll out its services and accelerate its operations by forging **partnerships designed to enrich its data catalogue**, by enabling the **development of innovative, high-impact uses**, and by helping to make France a **leader in the use of health data at the European and international level.**



Stéphanie Combes
*Director of the
Health Data Hub*

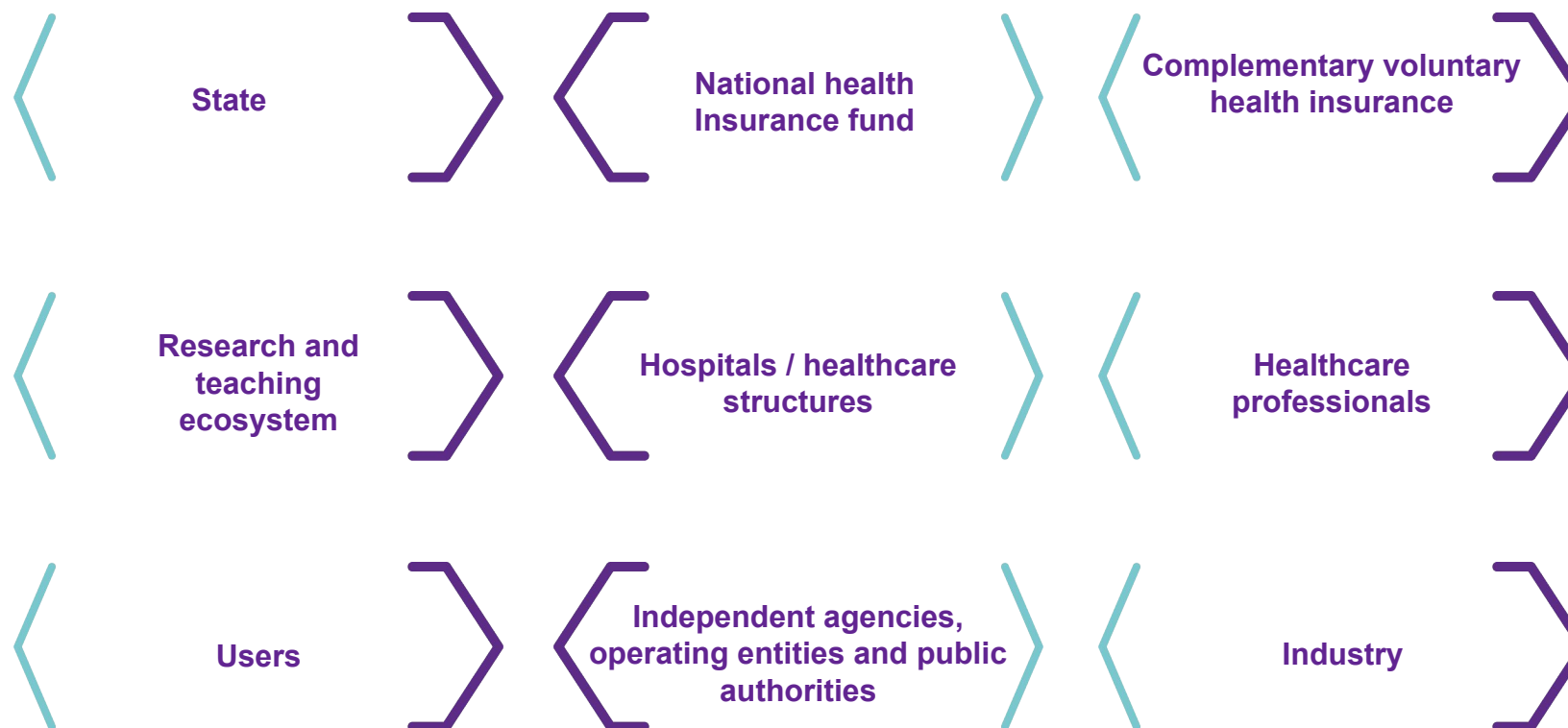
The Health Data Hub is a health data platform within the wider ehealth ecosystem

Health + medico-social + social dimensions



The Health Data Hub is a public interest group...

One of its missions is to federate the French health data ecosystem.
56 actors are part of its General Assembly.



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

... which aims to simplify access to data

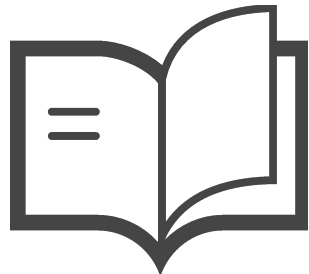
Established in December 2019, the Health Data Hub guarantees easy, unified, transparent and secure access to health data with a view to improving the quality of healthcare and support for patients. Its mission is divided into four main activities



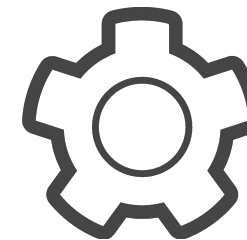
A **unique gateway**



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

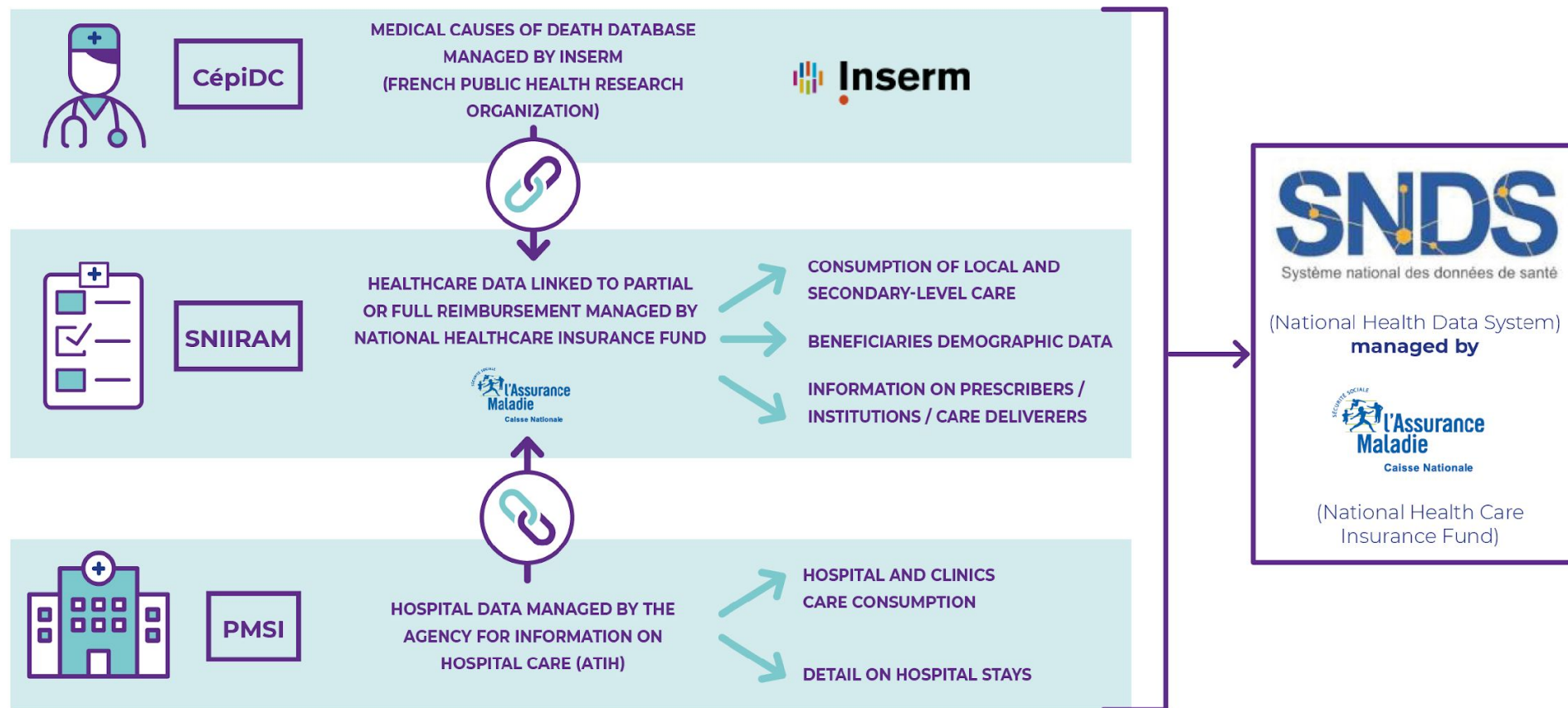


A documented, progressively
developed **data catalogue**



A **range of tools** to promote the
networking of key players in the
sector

In 2016, the creation of the SNDS is a major step forward for health research



The 2019 Health Law expanded the SNDS from a single database to a system of databases

1 The expanded SNDS includes clinical data: data from registers, research cohorts, hospital data warehouses, etc.

2 The catalogue centralises a collection of pseudonymized databases

- ❖ The historical SNDS has been renamed the "**main database**".
- ❖ The catalogue is a **collection of databases** derived from the SNDS, which the **Health Data Hub is authorised to make available, in addition to the main database**.
- ❖ The catalogue will bring together **copies of mostly pre-existing databases** that are deemed most relevant to research and innovation. A Strategic Committee led by the French government will establish the main guidelines for this catalogue, which is intended to be scalable. **These databases are enriched with data from the main database whenever possible**.
- ❖ This collection is developed in a progressive and iterative manner, in partnership with data holders
- ❖ The law provides for **better visibility of the common repositories** of the entire ecosystem, and the **harmonization and facilitation of access rules** with the establishment of the Health Data Hub as a secure service enabling access to the catalogue.

Examples of databases considered for inclusion in the HDH catalogue

Historical SNDS

OSCOUR (Coordinated Monitoring of Emergency Departments) database

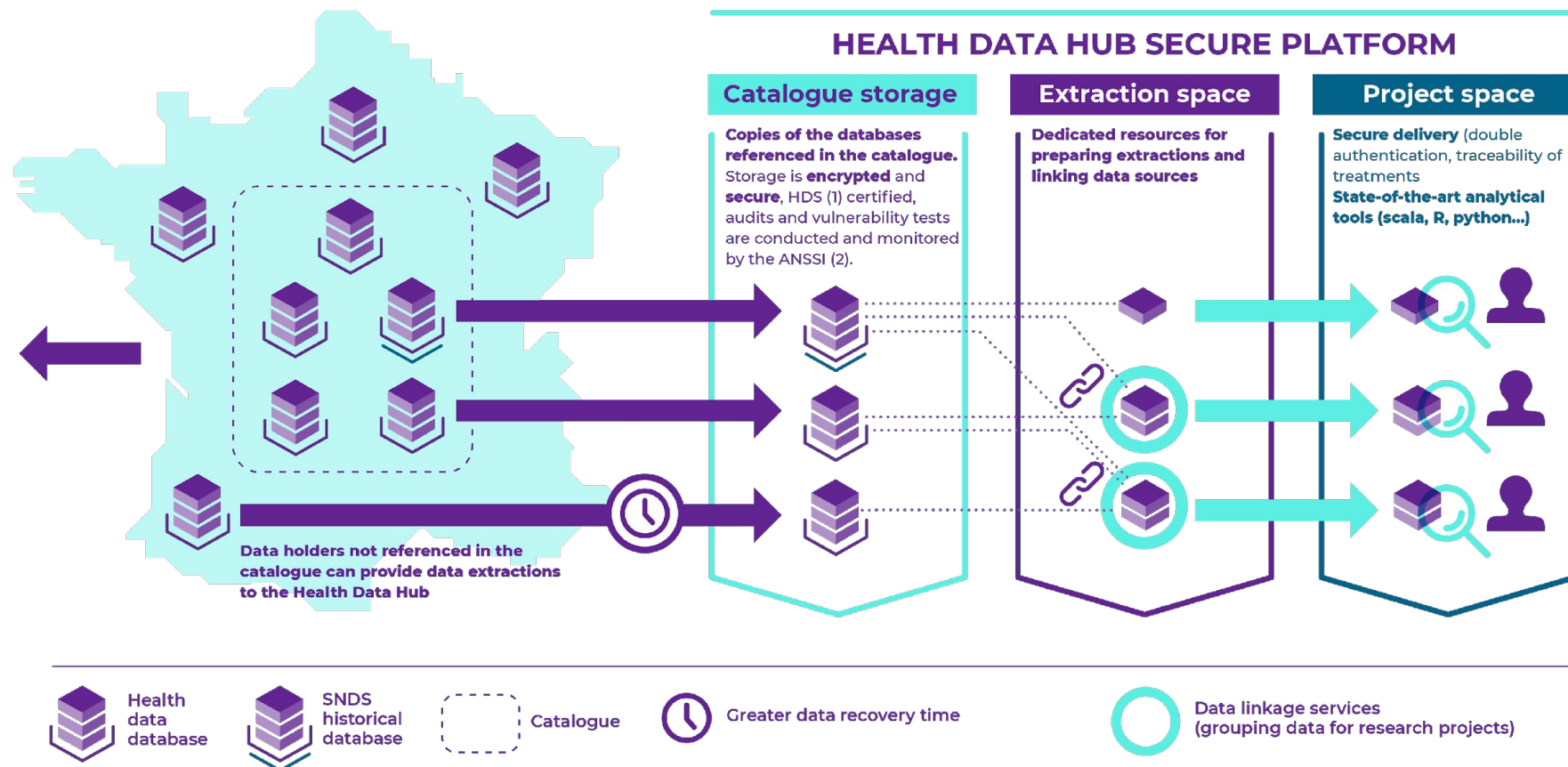
Registre France Greffe de Moëlle (French Bone Marrow Registry)

Cohort of sarcoma patients

I-share cohort

The Health Data Hub is authorised to make data available from the expanded SNDS (National Health Data System)

Health data catalogue: making data available



The functioning of the Health Data Hub is based on a precise legal architecture

Prerequisites must be fulfilled before the complete implementation of the catalogue in 2021

- 1 Publication of the **SNDS Decree (June 2021)**
- 2 Meeting of the **Strategic Committee defining the main priorities for the SNDS (June 2021)**
- 3 Opinion of the **CNIL (French Data Protection Authority) concerning the list of databases registered in the catalogue (August 2021)**
- 4 Publication of the **Order listing the databases (September 2021)**
- 5 **Signature of agreements** with data holders (ongoing)



Access to data from the catalogue is always restricted:

- ❖ **Prior authorization by the CNIL** is needed for the project leader, and requires an opinion from the CESREES (Ethics and Scientific Committee for Health Research, Studies & Evaluations)
- ❖ Each **study** and some of the **results** of the projects are **published on the Health Data Hub website**
- ❖ The **right to object** is granted to any citizen who submits such a request

Why is the HDH a "game changer" for health research?

Cross-referencing of sources

Starting patients' medical care as soon as possible



E.g. development of a tool to alert professionals via connected medical devices, enhanced with medical and administrative data annotations

Offering patients the best long-term treatments



E.g. comparison of medical care using cohort data enhanced with medical and administrative data when clinical trials are not possible

Elastic storage and computing capacities

Supporting healthcare professionals in an increasingly complex clinical context



E.g. characterization of alerts in case of drug interactions via the main SNDS database

Saving medical time, improving screening, and reducing diagnosis times



E.g. development of a diagnostic tool for 10,000 prostate MRI scans

Key figures

PROJECT SUPPORT

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

27 pilot projects and **1** Covid projects supported

+ with **33** partnerships under discussion with data holders **5**

SETTING UP THE ORGANISATION IN 2020

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

6 projects on the platform

25 databases expected for the first version of the catalogue

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)

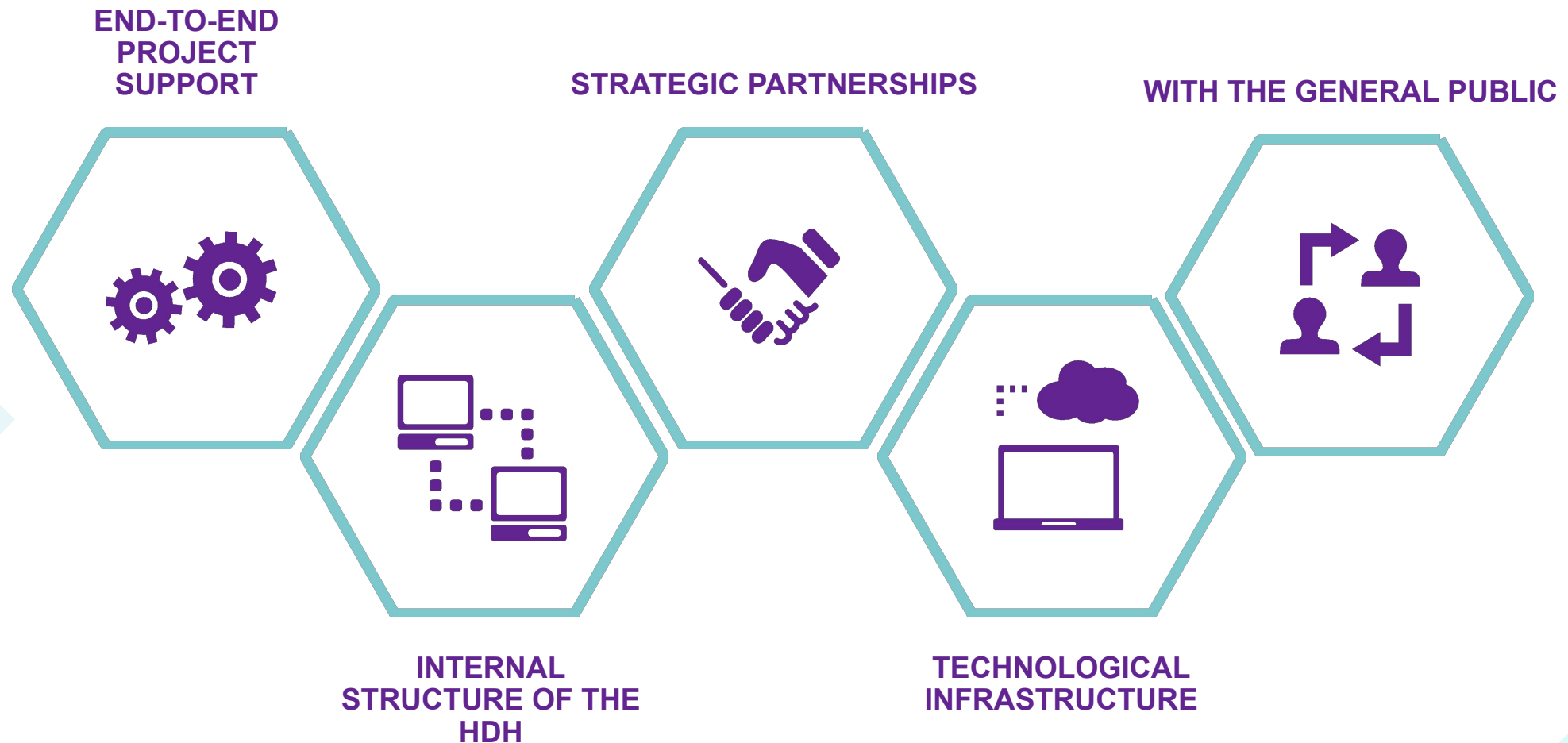
438 participants in the « AI4Health » Winter School

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

5 priorities for the Health Data Hub in 2021, to be implemented with its partners

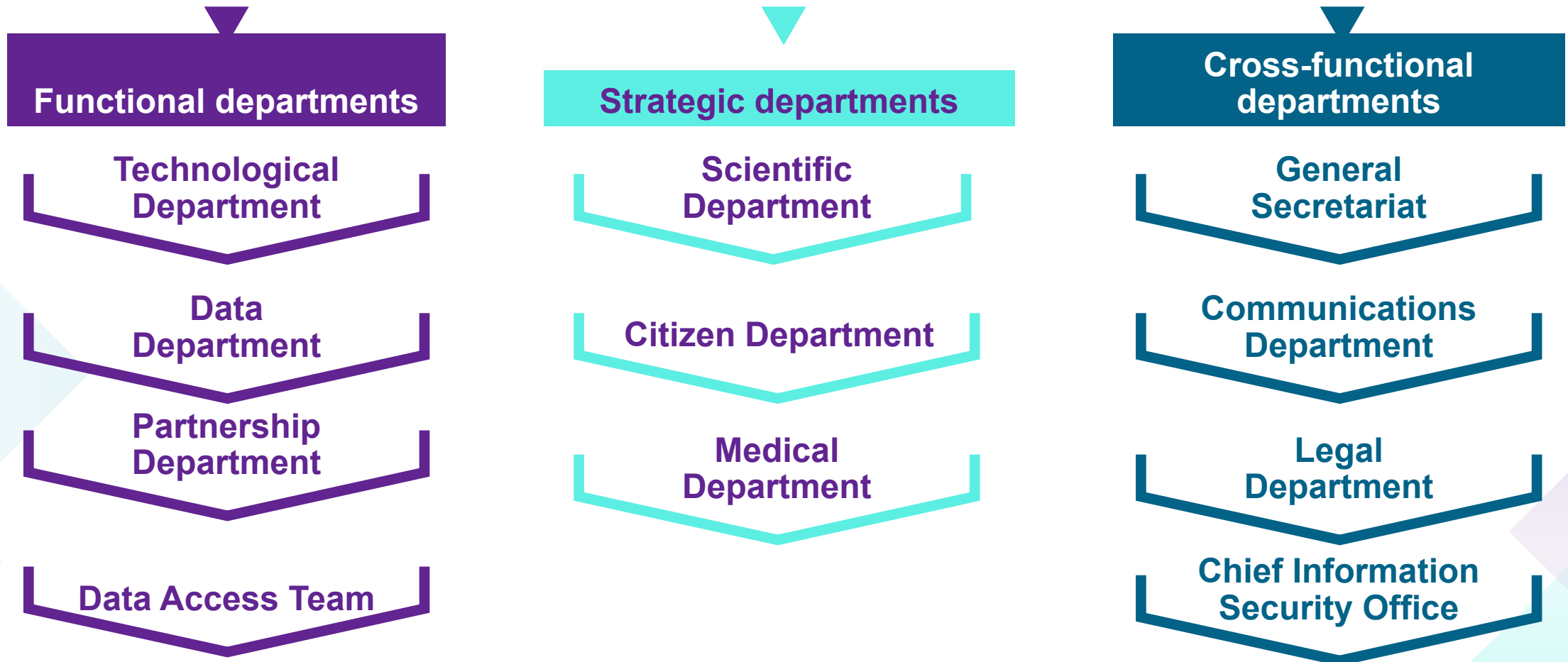




Challenges for HDH departments in 2021

Overview of the HDH's organisational structure

Executive Management Committee



The Data Access Team



When a project leader wishes to use data that are not completely anonymous, the French Data Protection Authority (**CNIL**) must authorize access to this data. The CNIL is the only body empowered to issue this type of authorisation.

Applications should be submitted to the Health Data Hub, which is the **single point of entry for requests for access to health data**. The HDH is tasked with ensuring that requests are processed in accordance with the regulatory framework, within the time frames defined by the law.

In 2020, **495 project leaders** applied through the Single Gateway for an authorisation from the CNIL.

As part of this process, the Health Data Hub also provides for the Secretariat of the **Ethics and Scientific Committee for Health Research, Studies & Evaluations** (CESREES).

This committee is responsible for **examining applications and, in particular, for assessing** whether the purpose of the proposed study corresponds to the scope of the data to which access is requested, the methodology presented, its ethical relevance, its scientific character, and the public interest of the project's objectives, where applicable. The Health Data Hub sends these applications to the CESREES within 7 days. **CESREES then has one month to issue its opinion.**

The Health Data Hub then forwards the file to the CNIL, which has a period of two months, renewable once, to provide a response.

To raise awareness of this system, the Data Access Team gives regular presentations to key players in the ecosystem.

The Data Access Team



The creation of a **body of documentation** to assist project holders with these procedures was launched in 2020 and will continue in 2021.

Questions of any nature, from details of the scope of the data requested to practical details about the information to be sent to the persons concerned, are encouraged in order to improve the understanding of these procedures which are often considered complex by participants at first.

In addition to designing educational resources for users, training is also provided in order to raise awareness, at different levels, of the **regulatory procedures and the functioning of the system as a whole**.

In addition, an **application management tool** will be implemented in 2021 to facilitate interactions between all stakeholders: project leaders, CESREES members and experts, its Secretariat, and data custodians or their governance bodies, where applicable. This tool will **enable the effective implementation of requirements for the transparency and sharing of the results of the processing operations carried out using the SNDS**.

In 2021, the Health Data Hub will also be collaborating with the CNIL on the development of several **simplified procedures**.



Valerie Edel
Director of the Data Access Team

The Partnership Department



Today, the Partnership Department is supporting **27 pilot projects**, 20 of which were selected following calls for projects. The latest project – to improve the user experience of the healthcare system – was organised in 2020 with Olivier Clatz, director of a major programme entitled "Improving diagnoses with artificial intelligence".

Eight projects have already been authorised by the CNIL. This authorisation is essential as it is a prerequisite for effective access to the data required to conduct these projects. Based on health data that are not name-linked, these projects aim to **improve the quality of healthcare in different fields**. One example consists of comparing the care provided for patients with rare cancers, and another concerns the development of a tool capable of automatically detecting tumours in mammographic images. To these ends, the Health Data Hub's technology platform will enable the secure analysis of data, while respecting patients' rights.

For the development of digital solutions to improve the quality of healthcare in the future, we are fortunate in France to have access to a pool of health data that is probably unparalleled in the world. This precious resource is the fruit of the remarkable work carried out by the people **in the French Health Insurance System (L'Assurance Maladie), health institutions and research bodies, among others**, who design databases, collect and compile data, and strive on a daily basis to ensure their quality and the scientific relevance of the uses to which they are put. One of the Health Data Hub's founding missions was to **compile a catalogue of "priority" health data** in order to facilitate its cross-referencing and provision, in a de-identified format, to the leaders of public-interest projects.

The Partnership Department



This catalogue is being developed in a progressive manner, in partnership with data collectors. It will be implemented by a decree issued following a decision of the SNDS Strategic Committee, which defines the broad guidelines for its composition, and following the opinion of the CNIL. In 2021, the aim is to **implement an initial version of this catalogue**, whose forerunner was the "entrepôt Covid-19" (COVID-19 warehouse) in 2020, developed following the Order of 11 July 2020 implemented in response to the national health emergency.

This initial collection of databases relating to the epidemic includes data from Santé Publique France on visits to accident and emergency departments, and data on hospitalisations of COVID-19 patients and their consumption of healthcare. These data originate from the French health insurance system, the ATIH (Technical Agency for Hospitalisation Information) and health institutions. In 2021, they will be enriched with data from the Si-dep screening and vaccination information system.

To implement a broader catalogue, **some twenty partners have been identified in the ecosystem of registers, national databases, cohorts, health institutions, health agencies**, etc. Some of these partnerships will also forge the first links between the Health Data Hub and **local or thematic initiatives**, including a pilot project involving several health institutions, the planned Unibase project with Unicancer and the cancer centres, and the partnerships with the Normandy region and the French National Institute of Health and Medical Research (INSERM) on the first pilot cohorts.

The implementation of these partnerships is dependent upon the completion of the ongoing projects on data sharing procedures, and especially on the issues of scientific and economic development.



Carole Requillard
*Director of the Partnership
Department*



The Health Data Hub's data experts contribute to the service offering in two main ways: by providing support to partners thanks to their highly specialised expertise, and by operating parts of the technology platform.

For example, engineers will be responsible for **matching and targeting data from the catalogue**, in order to make it available to project leaders who wish to process data on the Health Data Hub's technology platform. The implementation of this service will be a major challenge for 2021, as the data catalogue will contain highly diverse types of data such as images, management data, medical reports, etc., with variable levels of documentation and standardisation.

In partnership with key players in the system, such as the **CNAM** (National Health Insurance Fund) and the **CNAV** (National Old-Age Insurance Fund), the aim will be to **design a solution that ensures the provision of linked data sets in a matter of months**. This solution will need to be accompanied by an **efficient metadata catalogue service** in order to assess the feasibility of projects before requesting access to the data.

The Health Data Hub, which won a call for projects issued by the European EH DEN network in July 2020, converts medical and administrative data from the "historical" SNDS into a format that is destined to become an international benchmark. **OMOP-CDM** (Observational medical outcomes partnership – Common Data Model) will enhance database interoperability and facilitate the sharing of tools, methods, studies and programs internationally. These activities will continue throughout 2021, and should enable the sharing of the database in the catalogue.



The potential of health data is now widely recognised, and many sources of data are available; however, simply accessing health data is not sufficient to guarantee fast and reliable results. **Together with its partners, the Health Data Hub develops, provides and maintains a series of tools** to facilitate the handling of data.

Because understanding the organisation, biases, potentials, and limitations of data places considerable demands on each project leader, the Health Data Hub supports the pooling of knowledge by favouring **open source** formats whenever possible. In particular, promising opportunities associated with synthetic data are being explored in collaboration with the ANSM (French National Agency for the Safety of Medicines and Health Products) and INRIA (French Institute for Research in Science and Automation), which could pave the way towards data mining and algorithm development based on totally fictitious data, and also be used for training purposes.

2021 will see the pursuit and further development of these projects with a strategic focus on the **creation of a library of validated programs** to facilitate targeting in the "historical" SNDS database.

The sharing of knowledge is also fostered by the **coordination of a community**. To this end, a first data challenge was organised in 2020 with the French Pathology Society and a Major Challenge ("Grand Défi") programme entitled "Improving medical diagnosis with artificial intelligence". This data challenge set more than 500 competitors from 40 different countries the task of solving a machine-learning problem that required them to identify cancerous lesions on nearly 5,000 virtual slides presented for the occasion. On the strength of this successful experiment, the **Health Data Hub and the Grand Défi are hoping to repeat the data challenge in 2021**.



Maxime Kestemont
Chief Data Officer

Technological Department



Development work on the **Health Data Hub's technology platform** began in the spring of 2019, and it was **brought into service in April 2020**. This platform has been a major technological challenge given the need to meet stringent security requirements while also providing a software package capable of meeting the needs of innovative projects that in the past were sometimes considered impossible to conduct in France.

The security context of the Health Data Hub's technological platform was defined in a procedure developed in collaboration with the ANSSI (French National Agency for the Security of Information Systems) and the Senior Defence and Security Official for the Ministries for Social Affairs.

As part of the follow-up to certification, the security of the technological platform is tested and reviewed frequently through technical audits conducted by ANSSI-approved independent service providers.

The ANSSI and the CNIL are also involved in approving each new version of the technological platform, and **a new security audit was conducted by the ANSSI in February 2021**.

Recent opinions issued by various institutions also recognise the high level of security and data protection provided by the Health Data Hub platform. Examples include the French **Council of State**, in its Ordinance of 13 October 2020, and **DINUM** (The Interministerial Directorate for Digital Affairs) in its opinion of 11 November 2020 on the audit of the Health Data Hub project as a major government-sponsored digital project.

In line with the development of a strategic vision of a sovereign Cloud at the French and European level, the Health Data Hub has also launched several workstreams relating to **contractual changes, legal instructions, and ad hoc technological projects at the national and European levels**.

Technological Department



At the national level, the Health Data Hub's technological teams, in collaboration with DINUM, will continue to **identify sovereign cloud solutions that meet the specified functional and security requirements**.

At the European level, the Health Data Hub is participating in **the European Commission's joint initiative to design a common health data space**. This programme includes a work package relating to the federated technological infrastructure that will need to be used for exchanges of data. The Health Data Hub also supports the **French Gaia-X hub** initiative, which brings together French cloud users who are interested in the emergence of a sovereign solution.

At the end of 2020, the CNIL authorised the first eight Health Data Hub pilot projects, with many more authorisations expected in 2021.

To cope with the addition of these numerous projects to the technology platform during the course of 2021, the Health Data Hub plans to implement **a product strategy** in order to design the most suitable software package for uses of artificial intelligence in healthcare.

In addition to serving partner projects, the aim is also **to extend the opening up of the platform's services to the ecosystem**, and to enable the production launch of **between 50 and 100 projects per year** in the medium term. This scaling up requires the **industrialisation of processes** that have only recently been put in place by the Health Data Hub's technological team.



Louis Pery
Technological Director

The Scientific and Medical Department



The Health Data Hub's Scientific and Medical Department supports all the functional departments but also conducts its own projects with the aim of **contributing to training and raising awareness of issues** relating to the reuse of health data, **participating in the structuring of the European landscape**, and **identifying and contributing to the design of the most suitable service offering** in view of the diversity of the ecosystems with which the Health Data Hub interacts.

As part of its approach to awareness-raising and training, the Health Data Hub organised its AI4Health **winter school with three Interdisciplinary Institutes of Artificial Intelligence** (3IA): MIAI (Grenoble), 3IA Côte d'Azur (Nice) and Prairie (Paris), in collaboration with the Institute for Brain and Spinal Cord (ICM), the University of Paris and the French Medical Informatics Association (AIM).

This event brought together 438 participants from 32 different countries, with ten internationally renowned speakers and the support of some fifteen industrial partners. **This experiment will be repeated in 2021.**

Other events are also organised as part of more specific partnerships, such as a **joint seminar involving the Health Data Hub, the French National Academy of Medicine and the Massachusetts Institute of Technology (MIT)**, the **"Health Data and Collective Intelligence" symposium with the Ministry for Solidarity and Health**, and the **Grand Défi** programme entitled "Improving medical diagnoses with artificial intelligence", whose third edition is scheduled for 2021.

In 2020, the Health Data Hub created an **affiliate network** that now includes 167 researchers from 62 institutions and 50 different fields, ranging from the humanities and social sciences to medicine, computer science and engineering.

The Scientific and Medical Department



The Health Data Hub's participation in **PariSanté Campus**, a project aiming to bring together major players in the health data ecosystem on a shared site formerly occupied by Val de Grâce hospital, should help to boost this network.

In 2021, the Scientific Department is aiming to **create similar networks in both the industrial and health ecosystems**. These networks will develop links between actors and connect them to the Health Data Hub, thereby significantly increasing the opportunities for national and international collaborations. A **foundation project** is also being considered to facilitate the financing of the most promising projects.

To further these ambitions and develop strategic scientific partnerships, the Health Data Hub will be supported by a **Scientific Advisory Board** composed of recognised experts in research and medicine.

On the European scene, the Health Data Hub is maintaining its involvement in the **European Commission's joint initiative to create a European Data Space** in the health sector.

In particular, it is coordinating the French response by working alongside the following affiliated entities: **ANS, INSERM, Aix-Marseille Université, Université Toulouse III, HCL, and experts such as DNS, DREES, MESRI, DGE, CNAM, FHF and Santé Publique France.**



Emmanuel Bacry
Scientific Director



Hossein Roman Khonsari
Medical Director

Citizen Department



Health data can contribute to **building a more efficient, personalised and humanistic health system**. This entails the development of projects of general interest designed to improve people's health. Because health data are highly sensitive, the role of the Health Data Hub is also to **protect them by ensuring their security and protecting people's rights**.

Access to health data is only permitted in compliance with an ethical, transparent and data-protection-oriented approach involving a **National Scientific and Ethical Committee whose members include two representatives of health system users**, and the **French Data Protection Authority (CNIL)** – the independent authority responsible for the protection of personal data in France.

The list of projects is published on the Health Data Hub's website, a single gateway for the submission of data-processing applications, where the **results of these projects** will also be shared in compliance with competitive procedures.

As a stakeholder in data-access procedures, the Health Data Hub has also been assigned the legal responsibility to inform people and help them with their procedures. This has led to the **publication of clear and easily understandable information on its website**.

To help with the design of adapted and widely accessible content that is easy to read and understand, the Citizen Department is backed by **working groups which include representatives of users and citizens**, and it also conducts civic consultations.

Citizens' Department



These measures enable the department to gather people's views and identify their needs in relation to specific issues.

In 2021, as part of the European Commission's joint initiative to establish a common Health Data Space, **France is jointly managing a special research project on civic engagement.**

In this context, it is co-organising a large-scale consultation with Belgium, the results of which may be presented at an event organised by the digital health delegation in early 2022, during the French Presidency of the European Union.

Because citizens and users of the health system are not only interested in projects based on data, but also wish to be actors in the system, the Citizens' Department implements a **partnership strategy with key players such as France Assos Santé**, in order to determine the most effective ways of **working together**.

Priorities in 2021 will include determining how the Health Data Hub can **assist the associative ecosystem with the collection of data, the simple processing of the "historical" SNDS, and the establishment of links with other actors in the ecosystem.**



Yvanie Caillé
Director



Caroline Guillot
Deputy Director

Cross-functional departments and positions



Established on 1st December 2019, the Health Data Hub continues to grow steadily in 2021, and has set itself the target of **recruiting some 30 new staff**.

Access to health data is only permitted in **compliance with an ethical approach prioritising data protection and cybersecurity**.

On this last point, the technological platform provided for projects that wish to use it guarantees data hosting certified to "health data hosting" (HDS) standards and complying with the applicable legislation, including the decree on the security standard for the French National Health Data System (SNDS).

Compliance with the security requirements is ensured via a **cyber security governance system which involves the Senior Defence and Security Official for the Ministries for Social Affairs and the ANSSI** (French National Agency for the Security of Information Systems).

As far as individuals' rights are concerned, the Health Data Hub has been assigned a legal obligation **"to inform patients, and to promote and facilitate their rights, including the right to object"**.

This mission means disseminating information as broadly as possible to all audiences in order to explain the purpose of the Health Data Hub and its missions, while coordinating the ecosystem and promoting individuals' rights.

It is accompanied by an **ambitious technological project to enable the simplified exercise of individuals' rights via a simple form, which will be introduced in 2021, in partnership with the CNAM**.



Clarisse Manoussis
Secretary General



Christine Tendel
Director of Communications



Thomas Duong
Legal Director and Data Protection Officer



Members of the Health Data Hub

Members

Government



Health insurance funds



Complementary voluntary health insurance bodies



Members

Research and teaching



Healthcare institutions



Healthcare professionals



Members

Healthcare System Users



Agencies



Industrial companies



Contacts

Link to our website



LinkedIn



Our commitments
to civil society



Applicants



Our latest news



