WINTER SCHOOL ProMIS

3rd Edition – Year 2024

"From the European Initiatives to the Italian response strategies: institutional public health reforms through a holistic approach"

5-6-7/03/2024

TURIN - ITALY

Carlos Alonso Peña
Director of the Data Office Division

ProMIS
PROGRAMMA MATTONE INTERNAZIONALE SALUTE

GOBIERNO DE ESPAÑA
VICEPRESIDENCIA PRIMERA DEL GOBIERNO
MINISTERIO DE ASUNTOS ECONÓMICOS Y TRANSFORMACIÓN DIGITAL
SECRETARÍA DE ESTADO DE DIGITALIZACIÓN E INTELIGENCIA ARTIFICIAL

Oficina del Dato

1. **Data space: a place of value creation**
2. The Spanish National Health DataSpace (ENDS)

The Data Office Division is the unit in charge of streamlining the management, sharing and use of data throughout all productive sectors of the Spanish economy and society.

Oficina del **Dato**

Europe has among its strategic objectives the creation of a single European data market that boost AI. The Data Office deploys this objective in Spain, supporting the deployment of data spaces and the achievement of strategic digital sovereignty .

Industrial data spaces

Deployment of data spaces in Spain

Data spaces are the vehicle chosen in Spain to enhance the data economy, promote innovation and advance in the generation of shared value.

The objectives pursued with their implementation in Spain are:

1. **Improving the performance of strategic productive sectors**: Promote the generation of new solutions and business models based on data and Artificial Intelligence.
2. **Guarantee of strategic digital autonomy**: Ensure national capacity to provide and control critical technologies for digitalization.
3. **Promotion of the generation of quality data**: Ensure the quality of the data through its ethical and effective governance.
4. **Enhancement of public data**: Maximize the value of public data, promoting its accessibility and use.
5. **Promoting a fair and sustainable data economy**: Promote an equitable environment in the use of data and ensure the equitable distribution of benefits generated by data.

PART I. Concept of Data Spaces

1. Data space as a value generation ecosystem
2. Context and diagnosis
3. Characteristics of data spaces
4. Dimensions of data spaces
5. Agreements for data space operation: trust framework
6. Integral governance of interoperable data spaces
7. Organization and control over data spaces
8. Data space architectures
9. Actors and relevant technological solutions

Part ii. Strategic axes of deployment plan

1. Strategic objectives
2. Strategic axes for deployment of national data spaces
3. Objectives, axes, and measures

Part iii. Heuristic considerations for data space deployment

1. Creating deployment conditions
2. Data space conception
3. Considerations for utility of shared data
4. Interconnection among data spaces
5. Decalogue of requirements for data space development

	Support for strategic productive sectors	Guarantee of strategic digital autonomy	Promotion of the generation of quality data	Valuation of public data	Fair and sustainable Data Economy
A. Enabling elements of viability and sustainability					
A.1 Generation of market conditions.	X	X			X
A.2 Search for social and territorial cohesion	X				X
A.3 Community formation	X			X	X
A.4 Interoperability guarantee	X	X	X		X
A.5 Independence of the technological solution		X			X
A.6 Deployment of experimental laboratories	X	X			
B. Trustworthy data governance					
B.1 Establishment of governance frameworks	X		X		X
B.2 Security and confidentiality		X			X
B.3 Assessment of organizational maturity	X		X		
B.4 Improvement of the quality of the data exchanged	X		X		
C. Driving efforts in key sectors					
C.1 Targeted support for industrial sectors	X			X	
C.2 Response to social challenges					X
D. Determinant role of public administrations					
D.1 Involvement in sectorial deployment	X				X
D.2 Mobilization of the value of public data			X	X	
D.3. Value capture by public administrations			X	X	X
D.4 Public administrations as an example		X		X	X

Data spaces under the PRTR

Within NextGeneration Funds (C12.I1), the regulatory bases and call for the deployment of demonstrator centers and use cases of data spaces with a budget of 150 million euros have been published:

Access to the Aid Portal
Provision 969 of BOE no. 16, 2024

Within NextGeneration Funds (C14.I2), the regulatory bases and calls for "Last Mile" have been published, including up to 50 M for Tourism data spaces:

Aid Portal of the Ministry of Industry, Commerce and Tourism
Provision 26601 of BOE no. 310 of 2023

List of activities related to the data space deployment action plan

- Development of unique projects led by a group of Regions (RETECH) and municipalities (EDINT).
- Nominative subsidy to support the activities of the national Gaia-X Hub
- Support for the development of basic security and sovereignty technologies (INCIBE, UPM)
- Development of activities around the Language
- Competitive grants for the development of industrial data space (NextGeneration Funds)
- Public sector data spaces
- National health data space

Governance of interoperability

- Effective data governance at each participant in a data space is essential to adequately provide and consume the data being exchanged. The family of UNE standards around data are a fundamental piece of continuous improvement, evaluation and certification.
- Technical, semantic, organizational and legal interoperability must be ensured through adequate data space governance performance. There are still many unanswered questions (EU level) about the governance and regulation of data spaces.
- Interconnection between different data spaces is the long-term goal!!

1. Data space: a place of value creation
2. **The Spanish National Health DataSpace (ENDS)**

ENDS Mission and Vision

Catalyze research and innovation in health, through the reliable and ethical use of disruptive data processing technologies and Artificial Intelligence, within the framework of the Spanish Digital Health Strategy

01 National infrastructure for research and innovation and design of public policies in the field of health (**secondary use**) contemplated in the Digital Health Strategy and in the Plan to deploy NextGenerationEU funds (C18.i6 €100M)

02 Offering **differential value to participants**, either through the application of cutting-edge techniques (advanced AI, computing capacity) or through the quality and quantity of their data.

03 Obtaining the **trust of citizens** through strict safeguarding of confidentiality, privacy and monitoring of ethical considerations.

04 Even though its purposes do not include continuity of care (primary use), the synergies generated between both uses will facilitate **health performance**.

Information resources

The ENDS is built on a federated approach where the data is made available in a secure environment constituted according to the Use case to be studied.

01 Available information sources and resources must be incrementally federated to constitute a **single catalog of data services** accessible internally to stakeholders and externally to researchers (FAIR principles)

02 The data does not have to reside in the ENDS prior to its materialization within a specific **use case**. Even in that case, it is possible to use federated learning technologies that avoid data transfer.

03 The **quality of the data**, its homogenization/standardization, the use of terminology servers and its metadata are key aspects to pursue. The use of **privacy technologies (PET)** is essential given the sensitivity of the information processed.

04 The "Technical Office for Standardization and Quality of Health Data" defines interoperability standards and norms, as well as **pseudonymization and**

Target situation. Case of use

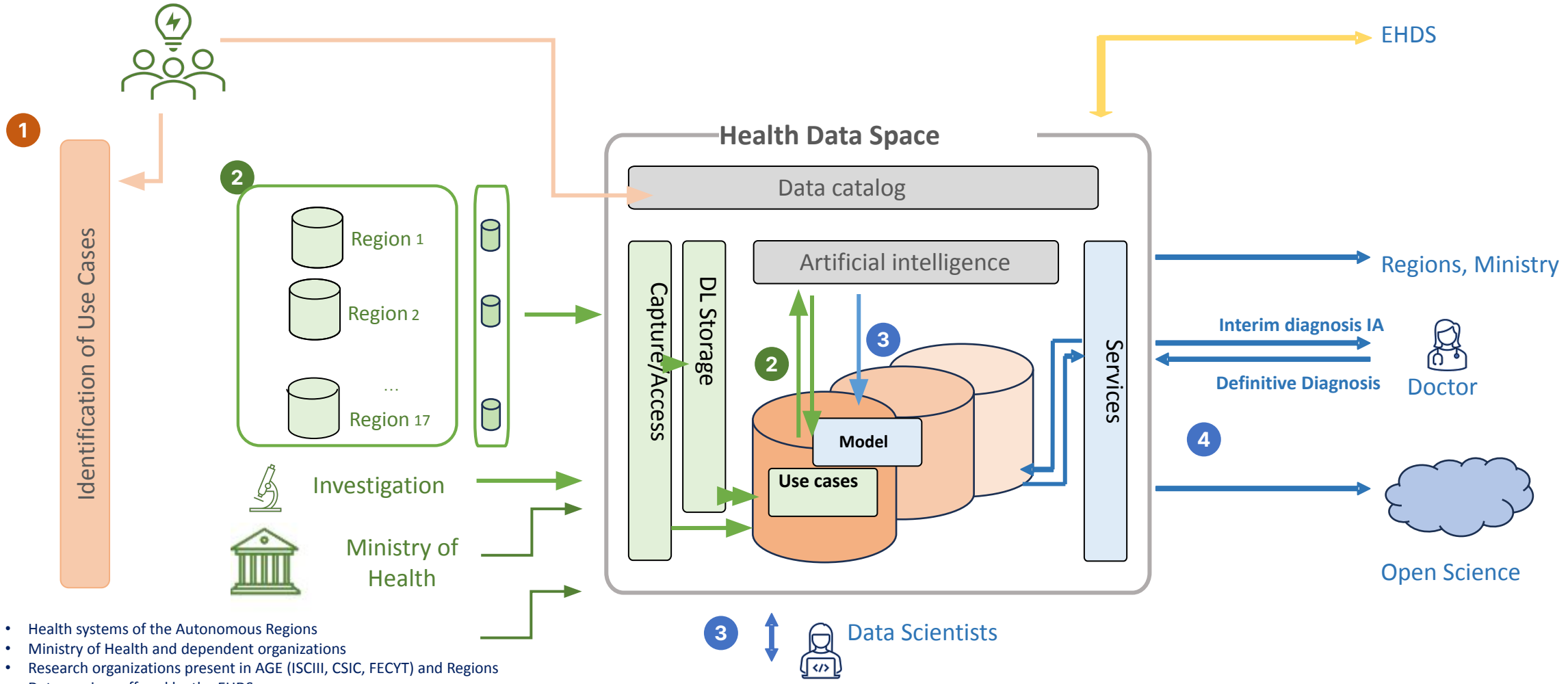
Every use case is developed in a differentiated Secure processing environment, deployed in a public cloud, with access to the tools and data required.

Applications for the development of use cases will be judged by an **Access Committee** from an ethical prism and contribution to the common good. Its results will be subject to validation by the appropriate scientific committee.

The research process will be governed by the principles of **open science**, prioritizing the sharing, not only of results, but of the research process itself, as soon as possible.

The execution environment must guarantee the correct **publication of results** (persistence of results, unique identifier of DOI sources, federation of CRIS research results, etc.).

General diagram

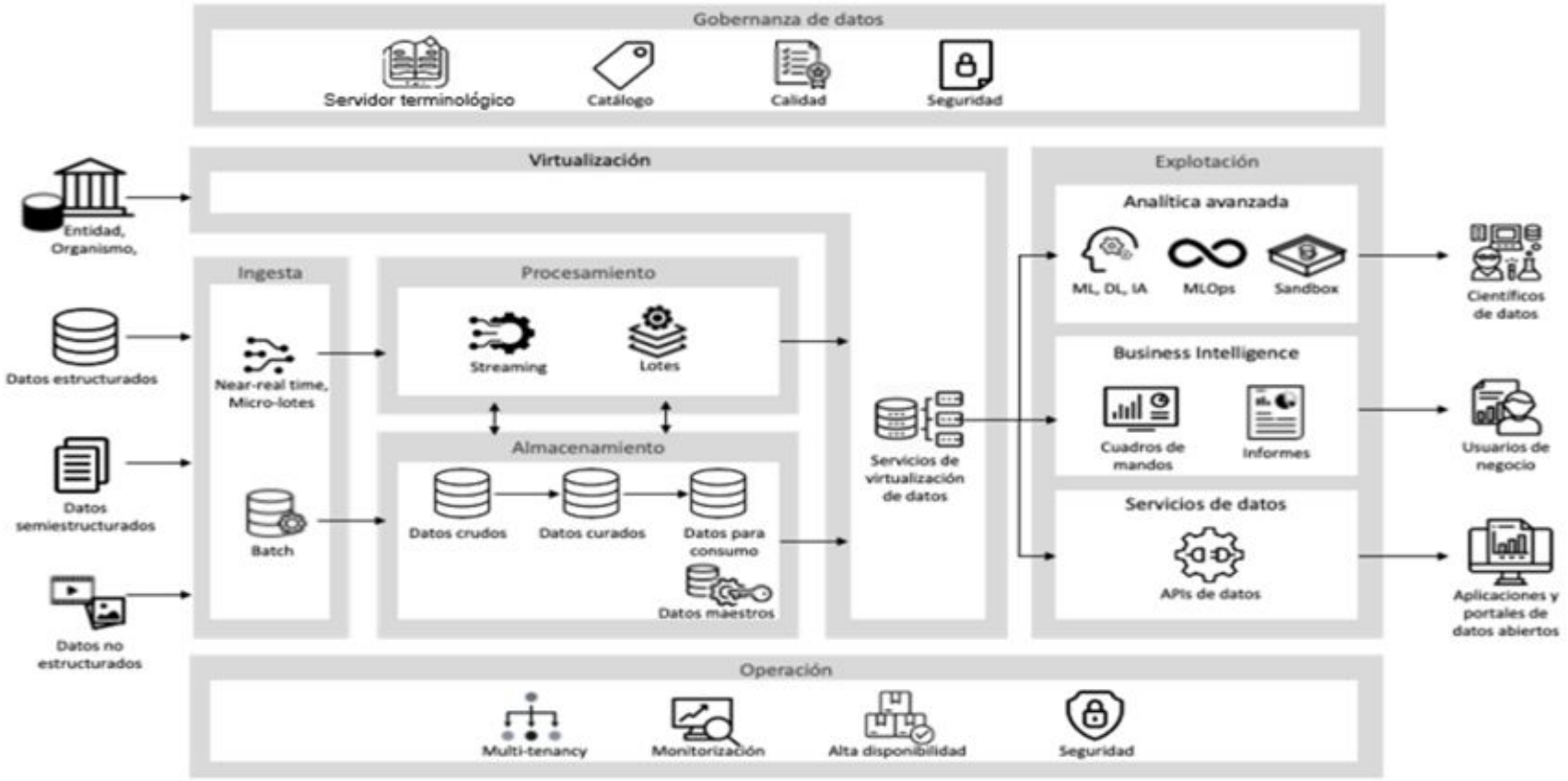


- Health systems of the Autonomous Regions
- Ministry of Health and dependent organizations
- Research organizations present in AGE (ISCIII, CSIC, FECYT) and Regions
- Data services offered by the EHDS.
- Other information (SS, INE, AEAT, MITECO, …)
- Non-health data offered by sectoral data spaces
- Real world information (open data, social networks)
- Services offered from data altruism schemes.
- Obtaining/purchasing data from the private healthcare sector

Target situation. Technical solution

The current choice of the technological solution (architecture, tools, standards, interoperability) determines the medium-term possibilities of the system.

It is important not to acquire infrastructure that may be underused, and **to guarantee the taking advantage of the public cloud** in terms of vertical and horizontal scalability.



Current situation. Governance

Governance of National Health Data Space development (monthly plenary meetings)

Working Group of the National Health Data Space (in plenary session), depends on the Digital Health Commission of the Interterritorial Council of the National Health System

In which all the autonomous regions, Ceuta, Melilla, INGESA, AEMPS, AEPD, Ministry of Health, SGAD, Data Office, MUFACE, MUGEJU, ISFAS participate

Three working subgroups for the development of the ENDS:

Governance Subgroup

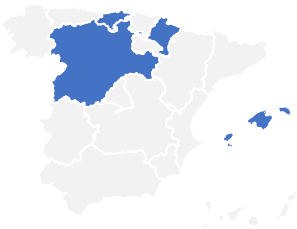
Propose, define and ensure the implementation of the policies, norms and standards that will govern the Health Data Space.

Participating: *Andalusia, Canary Islands, Catalonia, Galicia and AEPD. Also the DPOs of these Regions and the Ministry of Health*

Use Case Subgroup

Identify the use cases, both internal (National, Regions, Research Centers, etc.), and connection to the European Spaces.

Participating: *Balearic Islands, Cantabria, Castilla y León, CF Navarra and AEMPS*



Architecture, Standards and Interoperability Subgroup

Evaluate and define the technological models and processes that cover the needs raised, as well as the specific health standards and regulations for data management.

Participating: *Aragón, Valencian Community and Region of Murcia*

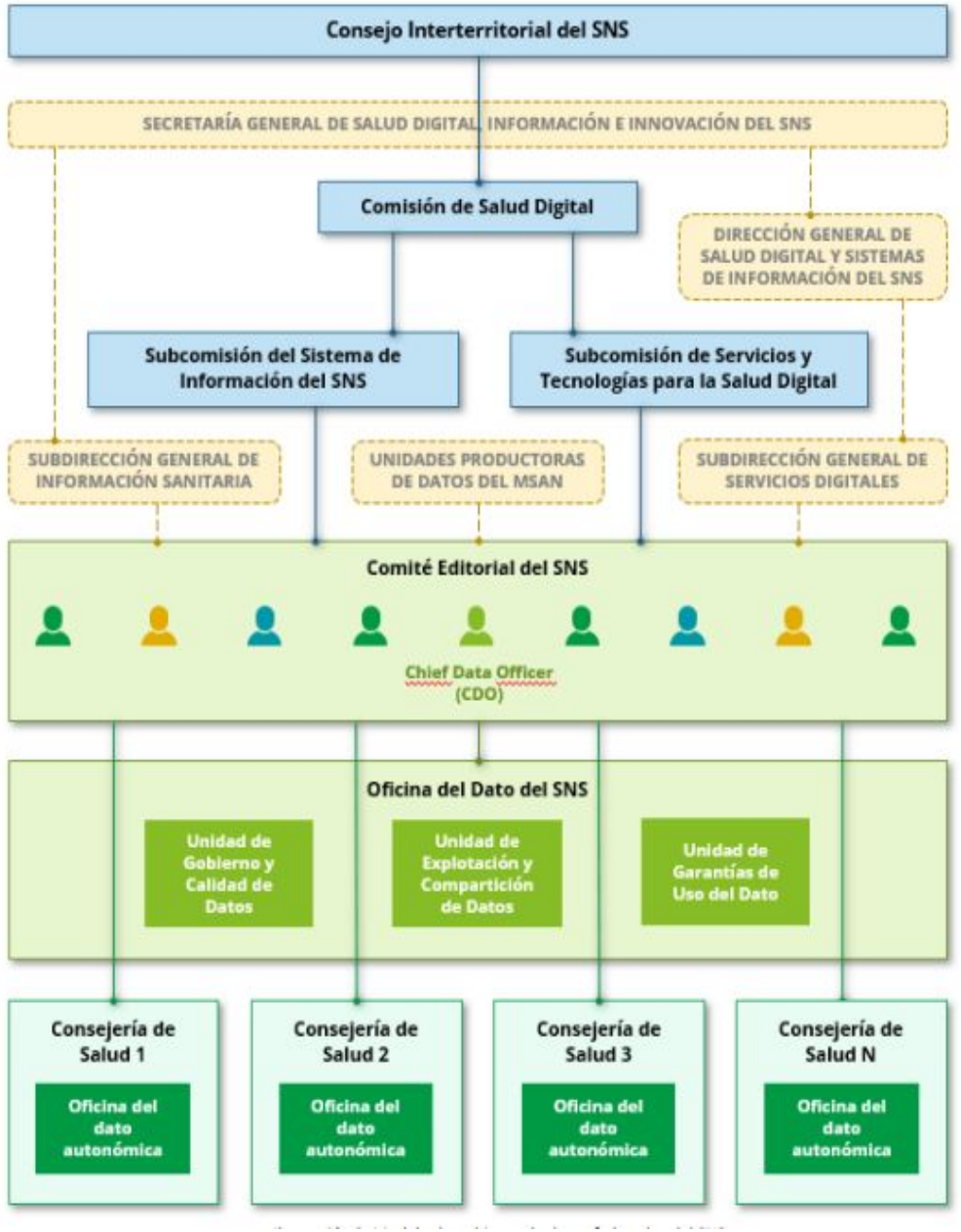
Current situation. Governance

Health Nacional System (SNS) data governance model

Coordination by the **Health Data Office of the National Health System**, following the guidelines and criteria agreed upon within the Interterritorial Council of the SNS.

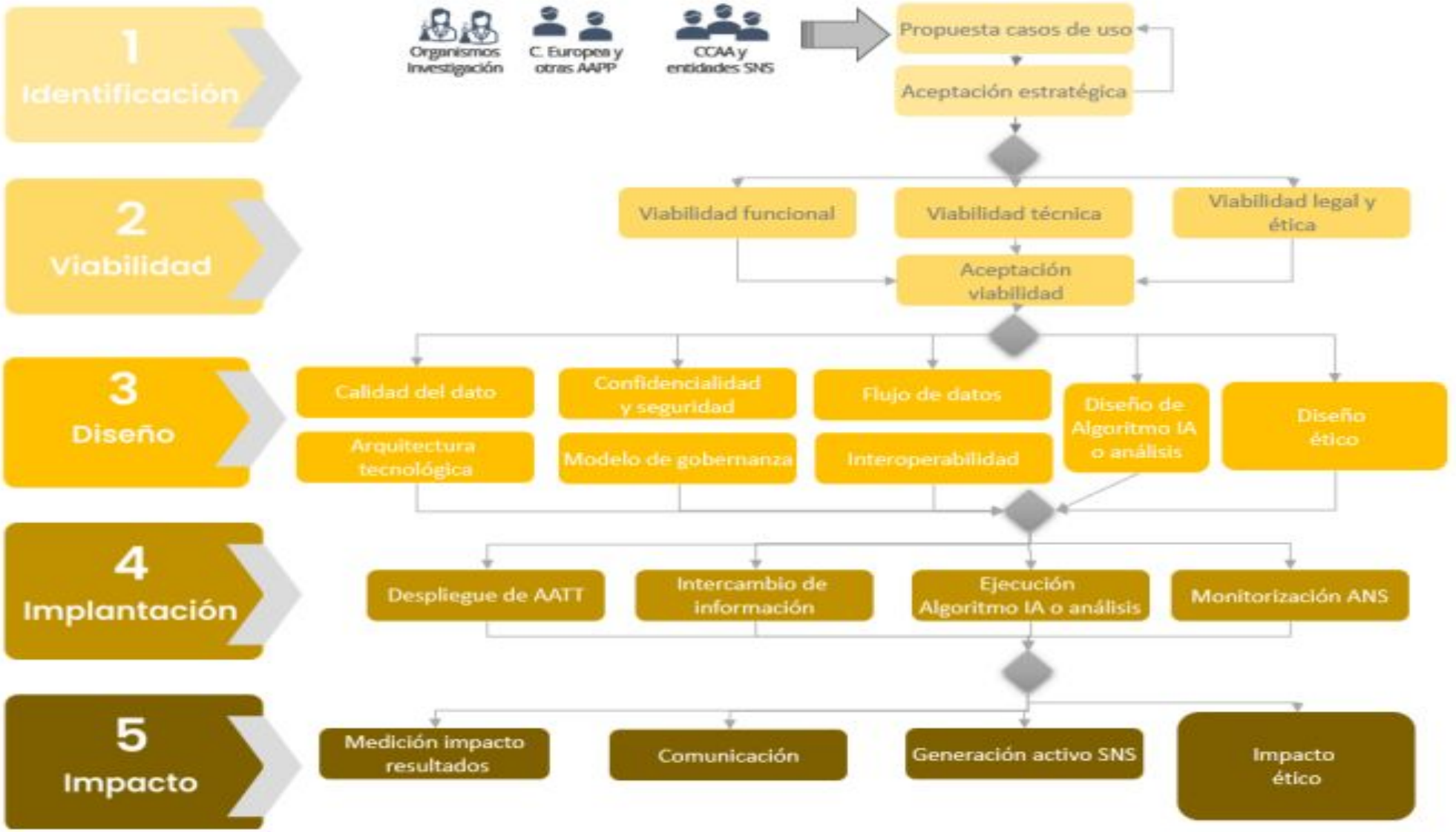
The activity to be carried out by the SNS Data Office will be approved, at the proposal of the Ministry of Health, by the governance bodies established in the Spanish Digital Health Strategy. The interterritorial commision of the SNS through the Digital Health Commission and its technical bodies.

The operational model of use case management has been defined supported by the UNE standards of governance, management and data quality.



Target situation. Case of use

Life cycle of the use case



Development of Use cases

It will start with the use cases of the 3 areas identified in the working groups

Area 1: Health-related data analytics for public health policies

Area 2: Monitoring the safety of medicinal products

Area 3: Deepening knowledge of illnesses

- Secure private infrastructure with high computational capabilities
- Larger and more diverse data sample that reduces the error of the models.
- Advanced análisis techniques: neural networks, image reading, machine learning, deep learning, natural language processing...
- Enriched data in Central Node: geopositioning, economic data, environmental data, social services, genomics, social network,...
- Communities of practice at national level

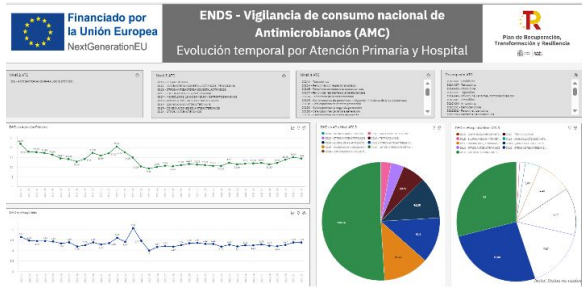
Use case : Medication Safety Monitoring

Describe the use of antibiotics, quantity, type of drugs used, indications of the prescriptions.
Temporal evolution (COVID effect, other virus waves), territorial differences and therapeutic adequacy with the recommendations of the guidelines.
Identify the characteristics of patients/pathologies that are associated with a choice different from that of the guidelines.

Phase 1 (target CID 284)

Objective : Obtain antibiotic consumption at the state level.
Necessary information : Antibiotic consumption at the state level.
Origin : **PRAN** (National Antibiotic Resistance Plan) Health

Product : Report on the use of different antibiotics in DHD (relative uses of the different groups, beta-lactams, macrolides, etc.) and the temporal evolution.



Phase 2 (to be developed)

Objective : Obtain antibiotic consumption in the SNS; Use rates adjusted by age and sex to be able to see variations in use between different Regions

Necessary information: Antibiotic consumption at the individual level together with data on the assisted population of each of the geographical units.

Origin : **BIFAP (AEMPS)** and data provided by the rest of the regions

Product : Report on the use of different antibiotics in DHD (relative uses of the different groups, beta-lactams, macrolides, etc.) in the different units (ZBS, ABS).

Phase 3 (to be developed)

Objective : Determine the use of antibiotics in different pathologies in non-hospitalized people and their suitability or not to the pharmacotherapeutic guidelines.

Necessary information: Antibiotic consumption at the individual and pathology level, together with data on the assisted population and pathology of each of the geographical units. Data are needed at the person and pathology level.

Origin : **BIFAP (AEMPS)** and data provided by the rest of the regions

Product : Descriptive report with APRR (Appropriateness of antibiotic prescribing) and generation of specific indicators for the selected pathologies

Development of Use cases

Use case : Deepen the study of diseases with greater prevalence, social relevance and impact

Determine what factors can influence decompensation in people diagnosed with COPD to detect high-risk population subgroups.
Assess the impact of the COVID epidemic on these variables in the years 2020-2021.

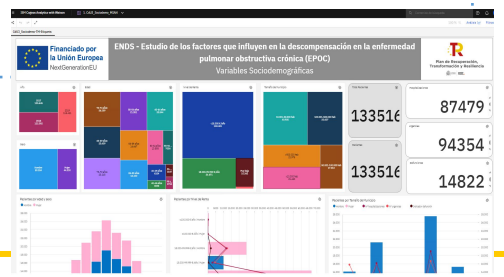
Phase 1 (target CID 284)

Objective : Determine the risk factors for mortality, emergency contacts and hospital admissions in patients with COPD and HF.

Necessary information : information on people with active health problems of COPD and HF in the BDCAP in the period 2017-2021

Origin : **BDCAP** (SNS Primary Care Clinical Database), **RAE-CMBD** (Specialized Care Registry-SNS Minimum Basic Data Set) and **BDPP** (SNS Protected Population Database).

Product : Descriptive analysis (2017-2021)
and predictive (2020-2021)



Phase 2 (to be developed)

Objective : Determine the risk factors for mortality, emergency contacts and hospital admissions in patients with COPD and HF.

Necessary information: People with active health problems of COPD and HF + geographic and environmental information from other ministries.

Origin : **BDCAP** (SNS Primary Care Clinical Database), **RAE-CMBD** (Specialized Care Registry-SNS Minimum Basic Data Set) and **BDPP** (SNS Protected Population Database). Air quality data in Spain (**MINTERD**). Climate records (**AEMET**)

Product : Descriptive analysis (2017-2021) and predictive analysis (2020-2021)

Target situation. Historical

2021

OCT
Constitution of working groups

2022

JUN
Reference Architecture Agreement for ENDS

DEC
Council of Ministers agreement for distribution criteria and distribution of 28M to regions

2023

MARCH
Sectoral Conference Agreement for the Digital Transformation of distribution to regions

JUN
Tender document for an ENDS (hw , sw , public cloud, services)

OCT
PRTR Addendum. Extension of implementation date to Dec/25. Includes new 466a objetive

DEC
software tender

CID #284 objective compliance

2024

FEB
Tender forecast open procedure for public cloud platform

MAY
Deployment of new use cases researchers

OCT
ENDS Platform

DEC
Work at CID 466a

2025

FEB
Regions Integration

DEC
CID #466a objective compliance – Development of 2 massive data analytics projects

Sharers. Synergies

The ENDS must be able to interoperate with the different national and European data initiatives.

Establishment of contact points with **health sector data spaces** within the framework of C12.I2 of the NextGenerationsUE funds

Search for synergies with the AI application use cases in the health sector developed in the **AI Value Chains and AI Missions program**.

Interoperability with European regulatory initiatives, **European Health Data Space (EHDS)** and its development pilots (HealthData@EU,HealthyCloud, EHDEN), as well as the European Open Science Cloud (EOSC).

Beyond the proposed framework

The ENDS must be able to interoperate with the different national and European data initiatives.

01 It is possible to think about **research aid** to be developed on the ENDS infrastructure, providing data or from the information present there.

02 Consider the **participation of startups** with innovative AI proposals. Employment of the GovTechLab initiative. EHDS establishes possible secondary uses (art.34) beyond clinical research, enabling access for small and medium-sized companies.

03 It is necessary to define the **sustainability of the project**, beyond the time frame of the existence of NestgenerationEU funds, as well as the possible participation of private actors.

04 In certain use cases, the system may require **supercomputing power (BSC)**, and a fluid relationship must be articulated with them due to capabilities and experience in

Grazie per il tuo tempo !!

Carlos Alonso Peña
Director of the Data Office Division