INFN EPIC Cloud

Building ISO/IEC 27001-certified Secure Cloud Regions for Trustworthy AI and Open Science

Barbara Martelli

ISGC26 - Taipei

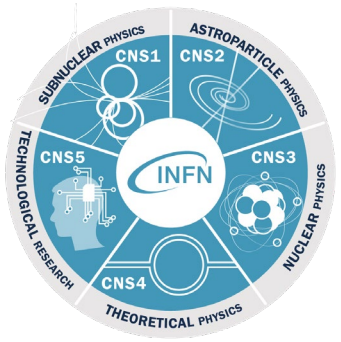
The starting point: INFN DataCloud

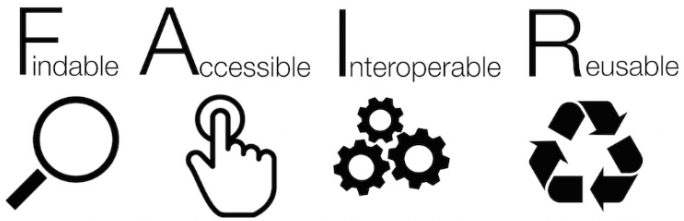
- The INFN Cloud portfolio, available via an **easy-to-use web interface** but also exploitable via **command-line interfaces**
- Based on **composable**, **scalable**, **open-source solutions** and can be easily extended either by the INFN Cloud support team or directly by end users
- Computing infrastructure distributed throughout Italy

https://www.cloud.infn.it/

Increasing security, privacy and compliance requirements

- From INFN life science and technology transfer communities
- From European life science projects
- From European regulations/best practices
  - GDPR
  - EHDS
  - AI Act
  - EOSC FAIR principles
  - NIS2





Need for computing platforms guaranteeing

confidentiality

integrity

availability

privacy

compliance

data sovereignty (as far as possible)

no vendor lock-in (as far as possible)

EPIC Cloud

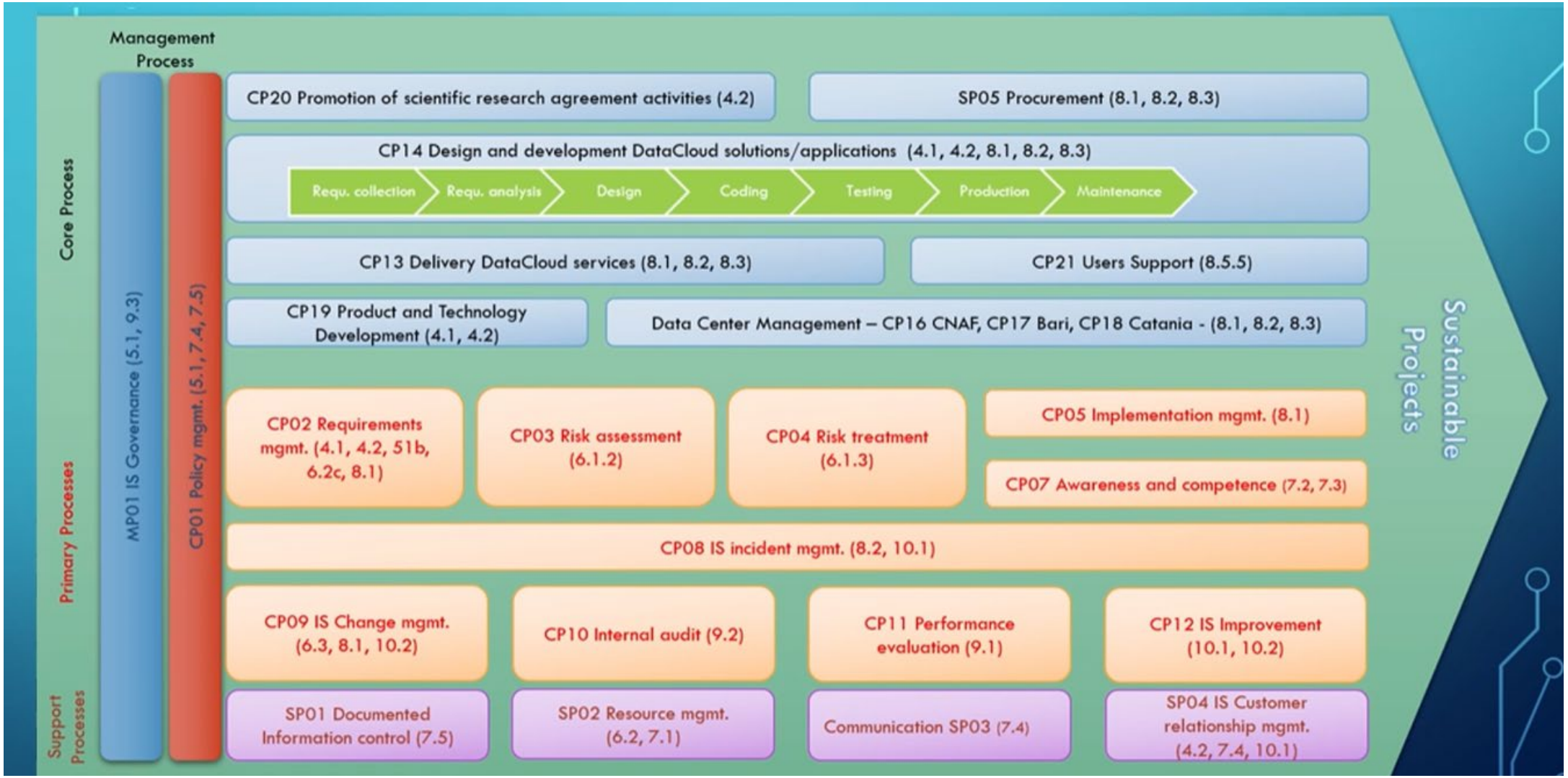
Enhanced **P**r**I**vacy and **C**ompliance Cloud

- Region of INFN DataCloud optimized for
  - medical research applications
  - technology transfer
  - public-private partnerships
  - all use cases with high confidentiality, integrity, availability and privacy requirements
- Certified ISO/IEC 27001 27017 27018 for
  - Community cloud
  - Provision of IaaS, PaaS, SaaS solutions
  - Co-development and co-design of advanced solutions
- Based on INFN DataCloud technologies

Link to INFN certificate

The Information Security Management System (ISMS) Process Reference Model

Based on ISO/IEC TS 27022
and M. Porter Value Chain Framework

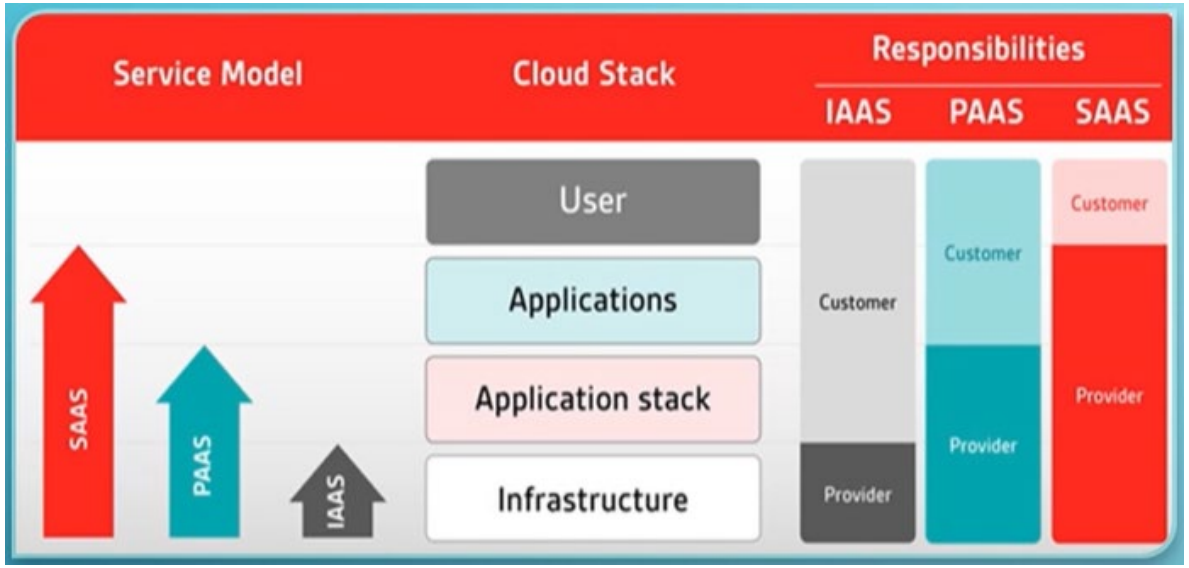


Organizational Transparency, Accountability, Effectiveness

Guaranteed by:

- Separation of Risk Assessement, Risk Treatment, Security Implementation
- Separation of ISMS management, Processes Evaluation and Improvement

- Clear Shared Responsibility Model defined in research agreements -> IaaS, PaaS, SaaS terms of use
- Data processing agreements to be signed before moving sensitive data to EPIC Cloud



Deming Cycle and Risk Assessment

Plan
(context, leadership, planning & support)

Establish the organization

- Define the scope
- Analyze the context and define the security requiremnts and objectives
- Define the Security Policy
- Determine roles and responsibilities
- Define the risk assessment approach

Do
(risk assessment & operations)

Risk Assessment

Control Selection

Control Implementation (Operations)

Check
(performance evaluation)

Monitoring KPI

Internal Audit
External Audit

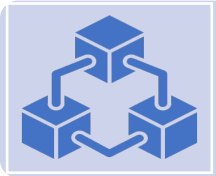
Act
(continual improvement)

Effectiveness

EPIC Cloud by numbers

Tens of active tenants

Hundred of users, several hundreds of human genomes, hundred of thousands of patients' datasets



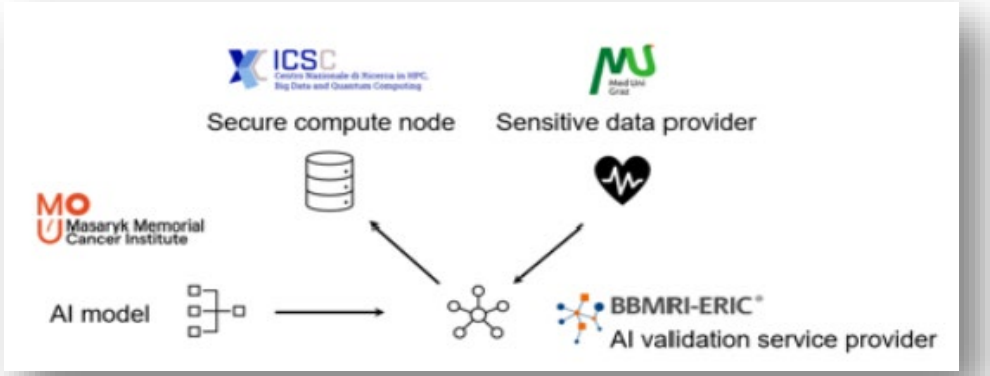
Site locations: Bologna, Bari and Catania

Resource available today:
Bologna: about 9 PB of HDD RAW, 1.5 PB of SSD RAW, about 12500 cores, 55 TB RAM, 62 GPU (6 A100, 36 H100, 20 L40S)
Bari: 3 PB of HDD (net), 1112 cores
Catania: 3 PB of HDD (net), 600 cores

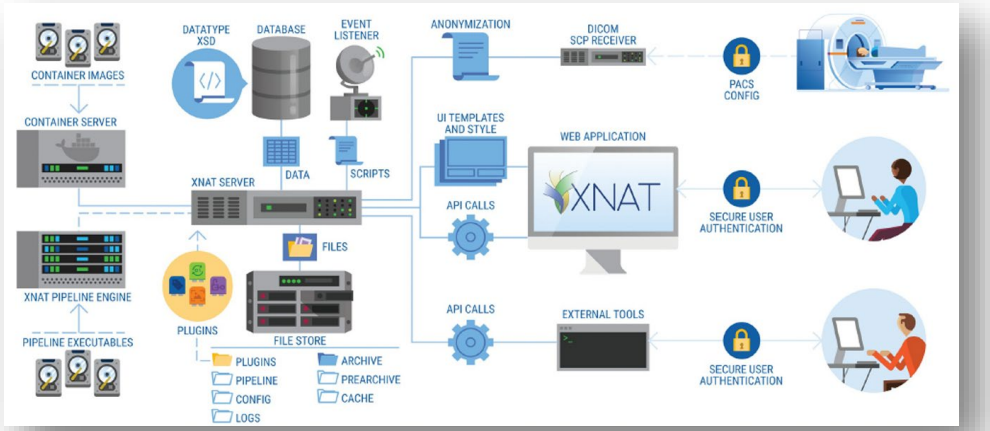
Ongoing expansion with further funds from external projects

Some examples of computing platforms deployed in EPIC Cloud

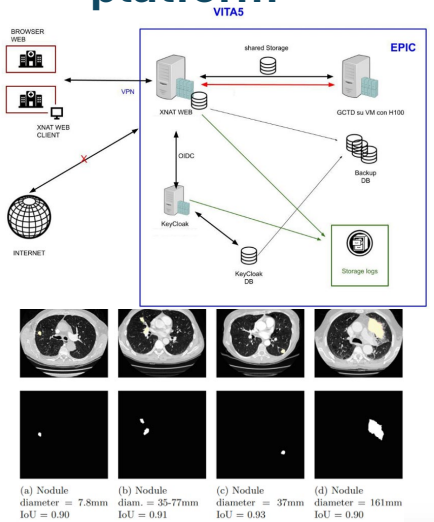
AI training and inference platform for biobanks



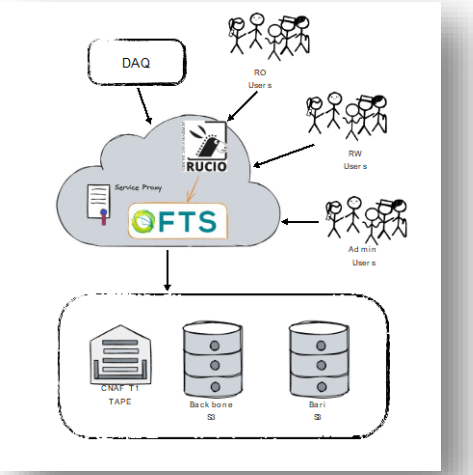
Medical imaging platform



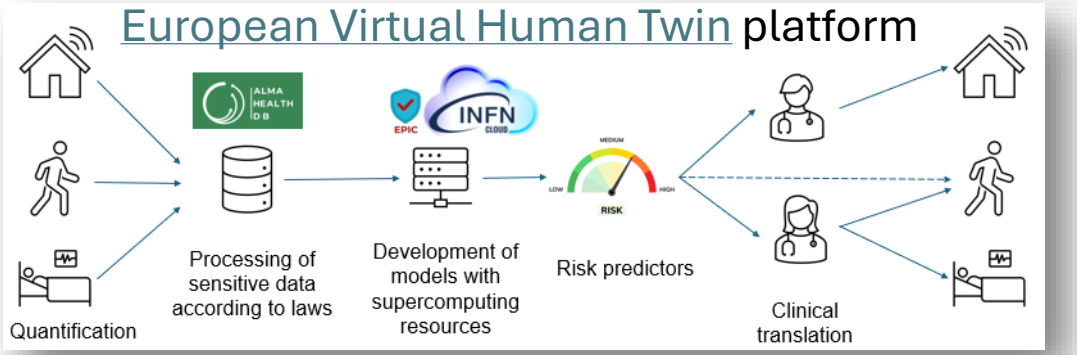
Virtual Imanging Trials platform



Computational Genomic data lake



European Virtual Human Twin platform



A process to extend EPIC Cloud capabilities through externally funded projects and Technology Transfer

bosco
bologna • sant'orsola • computational genomics

ICSC
Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing
(Spoke 0, 8)

BBMRI-ERIC®
biobanking for a healthier world

HARMONY ALLIANCE

HealthBigData

DARE
DIGITAL LIFELONG PREVENTION

ECHO TWIN

UNCAN.eu

COVIDSTAT

VTA

GOE
GENOME OF EUROPE

EUCAIM

ALLEANZA CONTRO IL CANCRO

NIH STRIDES
Accelerating biomedical research

elixir ITALY

AIFA
AGENZIA ITALIANA DEL FARMACO

- Express requirements, funds the hardware
- Co-design, co-develop new features
- Integrate the new feature in the DataCloud platform, extending its capabilities and fulfilling projects requirements

Conclusions

- INFN has developed a novel organizational approach to embed security, and privacy as integral components of INFN's organizational and operational ecosystem.
- Security and compliance are not treated as isolated functions but as value-generating activities embedded throughout the value chain
- Exploiting the PRM, INFN has developed an interoperable, dynamic and secure cloud that provides:
  - A Computational Genomics platform, aimed at collecting data from sequencing instruments
  - A platform for AI training and inference platform for biobanks
  - A medical imaging platform based on XNAT
  - A platform for the European Virtual Human Twin for preventive and precision medicine
  - A Virtual Image Trials (VIT) platform, with the goal of building a national center for VIT
  - Connectivity with EU initiatives such as ELIXIR, Genome of Europe, EOSC, and EHDS
- All of this in compliance with the regulatory framework of the European Health Data Spaces, including GDPR, the AI Act, and NIS2.