# EPIC Cloud
# a secure, GDPR-compliant, open-source cloud platform for life-science applications

Barbara Martelli, Davide Salomoni

*IBERGRID 2022 - Faro*

# Outline

- What is EPIC Cloud
- Motivations
  - Some life science projects and their requirements
- The ISO/IEC 27001 standard: what is an Information Security Management System (ISMS)
  - INFN implementation
- Future work: expansion of the Management System
- Summary and final thoughts

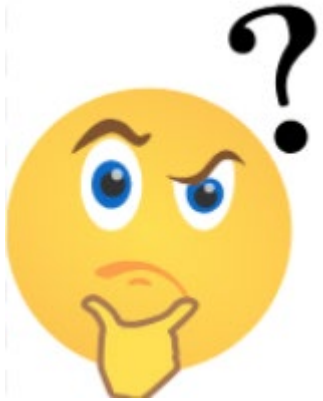
# What is EPIC Cloud (1/2)

| | |
|---|---|
| Enhanced PrIvacy and Compliance Cloud is an ISO certified cloud platform | A region of INFN Cloud with a certified Information Security Management System |
| EPIC Cloud offers an IaaS Community Cloud for the communities of | Biomedical and genomic researchers<br>Industrial researchers |

# What is EPIC Cloud (2/2)

- It is based on the same technologies of INFN Cloud (OpenStack, Indigo IAM, OneData), with various enhancements introduced to meet higher security and privacy standards. For example:
  - The IAM is federated with FreeIPA and Keycloak -> provides 2FA, integration with web services, SSH and VPN (OpenVPN)
  - Enhanced ONEDATA with more auditing functionalities
  - Network segregation between OpenStack tenants is guarantee by ACLs
  - At-rest and in-transit encryption
  - Standard shared responsibility model:
    - User manages data, applications, runtime, middleware and OS
    - EPIC manages networking, storage, servers, virtualization
  - Advanced logging and auditing services
    - centralized syslog managed applying the ***segregation of duties*** principle

# Motivations



## Why do I need to spend so much effort in deploying an ISMS?!

Explicit external stakeholders' requirement

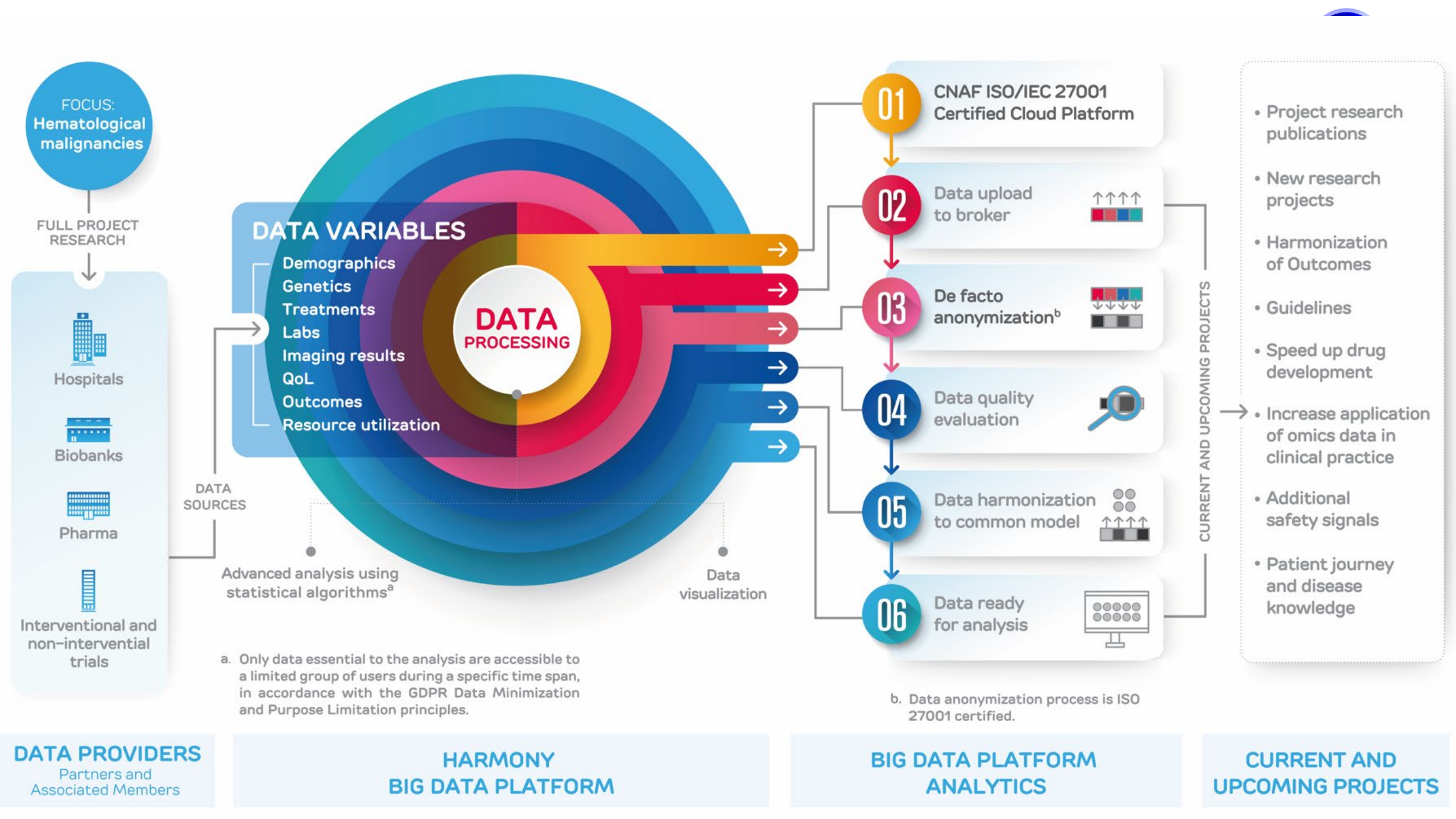
Need to guarantee compliance privacy and security laws and regulations

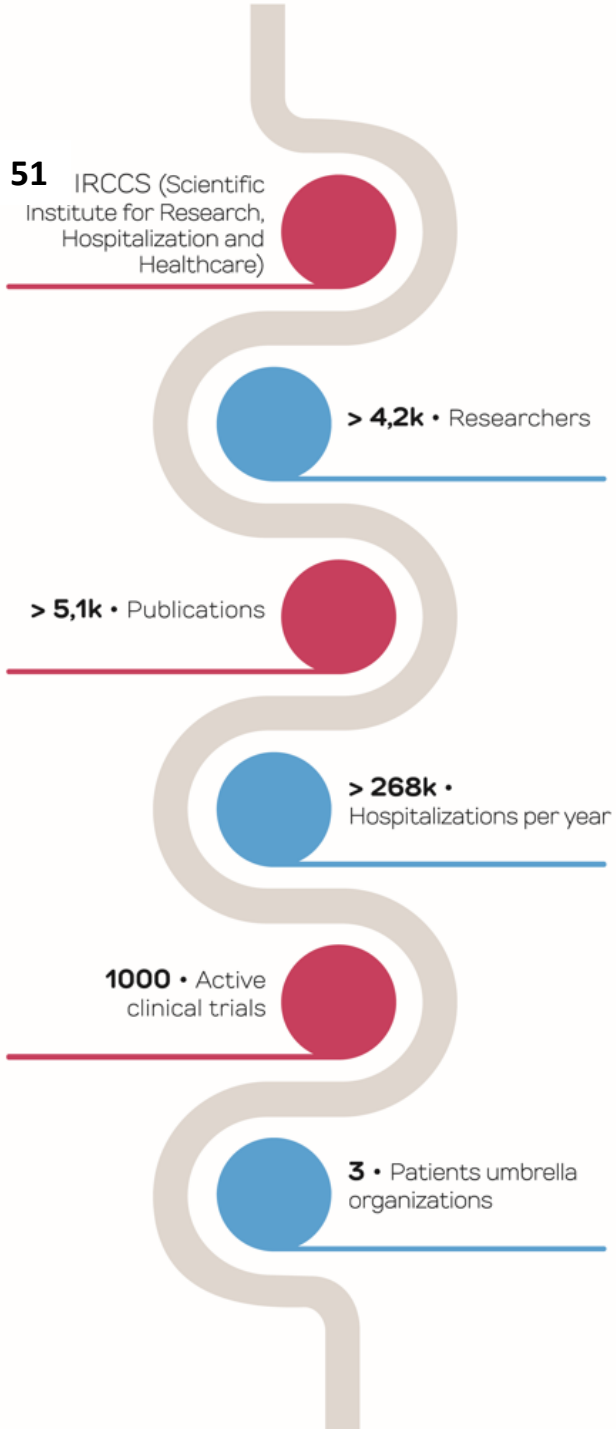
- Life science data -> GDPR and national privacy authorities' regulations
- Cloud is always considered a *critical infrastructure*

# Requirements coming from Harmony Alliance

- In 2016 CNAF decided to take part in the HARMONY ALLIANCE (Healthcare Alliance for Resourceful Medicine Offensive against Neoplasms in Hematology)
  - Harmony is a public-private partnership involving more than 100 organizations from 18 European countries, such as hospitals, universities, research institutes, medical associations, patient organizations, pharmaceutical companies and IT companies
  - Harmony is aimed at exploiting Big Data to develop more personalized treatments for blood cancer patients
  - Most of data analyzed in Harmony are genomic data -> they fit in the GDPR category of *particular data*
  - The project big data platform is located at INFN-CNAF
  - Mandatory requirement from the project to manage such a platform: being certified ISO/IEC 27001

100k patients' datasets


FOCUS:
Hematological malignancies
FULL PROJECT RESEARCH
Hospitals
Biobanks
Pharma
Interventional and non-interventional trials
DATA SOURCES
DATA VARIABLES
Demographics
Genetics
Treatments
Labs
Imaging results
QoL
Outcomes
Resource utilization
DATA PROCESSING
Advanced analysis using statistical algorithms[a]
Data visualization
01 CNAF ISO/IEC 27001 Certified Cloud Platform
02 Data upload to broker
03 De facto anonymization[b]
04 Data quality evaluation
05 Data harmonization to common model
06 Data ready for analysis
CURRENT AND UPCOMING PROJECTS
• Project research publications
• New research projects
• Harmonization of Outcomes
• Guidelines
• Speed up drug development
• Increase application of omics data in clinical practice
• Additional safety signals
• Patient journey and disease knowledge
a. Only data essential to the analysis are accessible to a limited group of users during a specific time span, in accordance with the GDPR Data Minimization and Purpose Limitation principles.
b. Data anonymization process is ISO 27001 certified.
DATA PROVIDERS
Partners and Associated Members
HARMONY BIG DATA PLATFORM
BIG DATA PLATFORM ANALYTICS
CURRENT AND UPCOMING PROJECTS




# Alliance Against Cancer/ Health Big Data

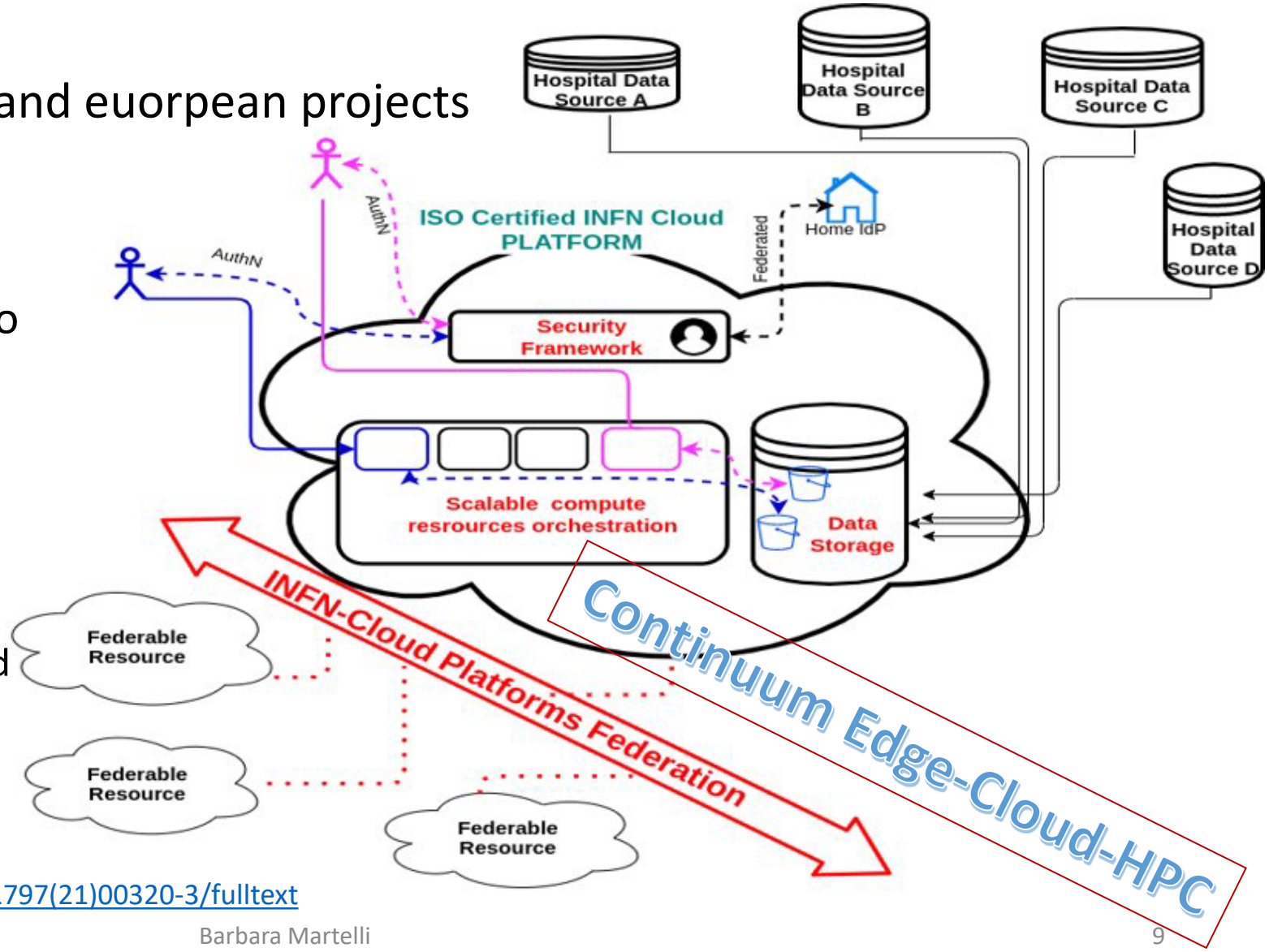
- A 10-year project founded by the Italian Health Ministry
- Goal: develop a federated cloud platform enabling the sharing of patients' data at national level (EHR, omics, clinical, epidemiology data)
- Almost all Italian Scientific Institutes for Research, Hospitalization and Healthcare involved

# EPIC Cloud enables the federated datalake for health-related use cases

Several other regional, national and euorpean projects

*(see Davide Salomoni's talk tomorrow)*

- In general we receive heterogeneous requirements form hospitals: someone need central repository, others need to store data locally
- Possible scenarios:
  1. Central harvesting of data collected remotely
  2. Edge-level anonymization and central ingestion and analysis of data
  3. Edge-level Feature Extraction and central ingestion and analysis of features
  4. Federated learning (training at local sites and algorithm publishing)




https://www.physicamedica.com/article/S1120-1797(21)00320-3/fulltext

# Why Lifescience communities required an ISO 27k certified system

In 2016 GDPR entered into force, it applies since 25 May 2018.

GDPR is about protecting natural persons with regard to the processing of personal data.

Genomic data like ones managed in our life science communities are personal data (fit in the Art.9 special categories of personal data) and are mostly impossible to be anonymized

GDPR shall always be applied

ISO/IEC 27001 is the main certification mechanism compliant with GDPR requirements

(Art. 43, 58, 63)

From the Controller side (hospitals), the fact that INFN is ISO certified is a way to **demonstrate that processing is performed in accordance with GDPR**

- *"Adherence to approved codes of conduct as referred to in Article 40 or approved certification mechanisms as referred to in Article 42 may be used as an element by which to demonstrate compliance with the obligations of the controller" (GDPR Art. 24 – Responsibility of the controller)*
- *"…the controller shall use only processors providing sufficient guarantees to implement appropriate technical and organizational measures in such a manner that processing will meet the requirements of this Regulation…" (*GDPR Art. 28 - Processor)
- *Adherence to […] an approved certification mechanism as referred to in Article 42 may be used as an element by which to demonstrate compliance with the requirements* set out in paragraph 1 of this Article*" (GDPR Art.32 – Security of processing)*

# …so, we got certified

First certification in 2017:

- Hosting of physical and virtual systems for biomedical data access and conservation, and management of data analytics applications targeted to biomedical and genomic research.

In 2021 we renewed the certification and extended it to also cover cloud services:

- Hosting of physical and virtual systems for biomedical data access and conservation, and management of data analytics applications targeted to biomedical and genomic research. Delivering cloud services in IaaS mode by applying ISO/IEC 27017:2015 and ISO/IEC 27018:2019 guidelines

Link to certificate

# ISO 27017 and ISO 27018

- In the years 2017-2020 we have improved our services by moving from a bare-metal platform to a cloud platform.
- Information security in cloud is subject of specific risks that need to be considered in order to provide a secure service: ISO 27017 and 27018 guidelines help organizations to cope with these risks
  - ISO 27017 Code of practice for information security controls for cloud services
  - ISO 27018 Code of practice for protection of Personally Identifiable Information (PII) in public clouds acting as PII processors

*If you offer cloud services and want to be ISO 27001 compliant, you must add 27017 and 27018 controls too*

# The ISO/IEC 27001 standard

*Requirements for establishing, implementing, maintaining and continually improving an Information Security Management System (ISMS)*

…but what is exactly an ISMS?

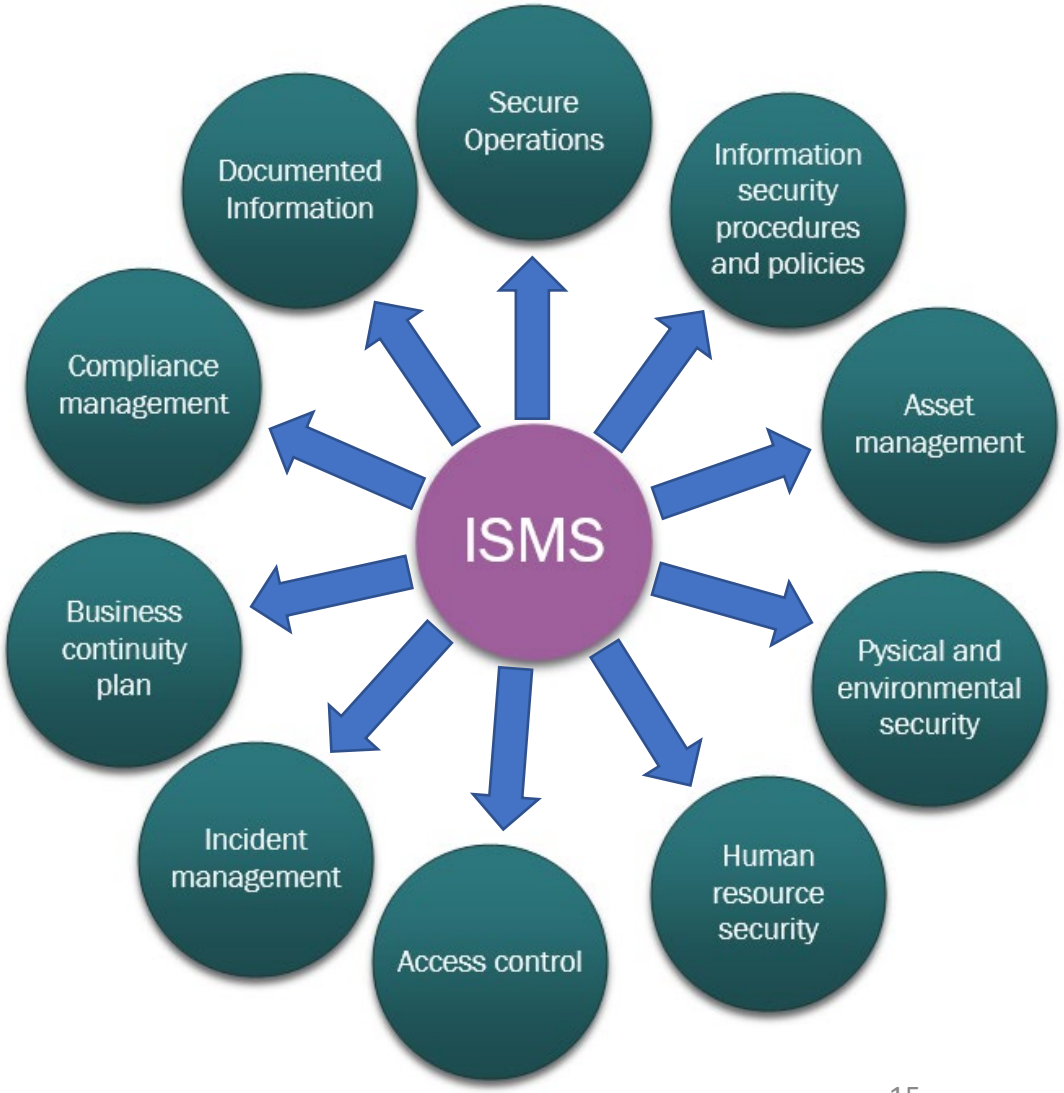
# Information Security Management System (ISMS)

- Information Security Management is about preserving the Confidentiality, Integrity and Availability (CIA) of information and associated information facilities (systems, services, infrastructure or physical locations)
- It ensures business continuity by minimizing business damage by preventing and reducing the impact of security incidents
- Other properties can also be involved, such as authenticity, accountability, non-repudiation, reliability and FAIRness
- The objectives of the ISMS are NOT fixed, they depend on the context and are defined by the organization

# ISMS: What's all about

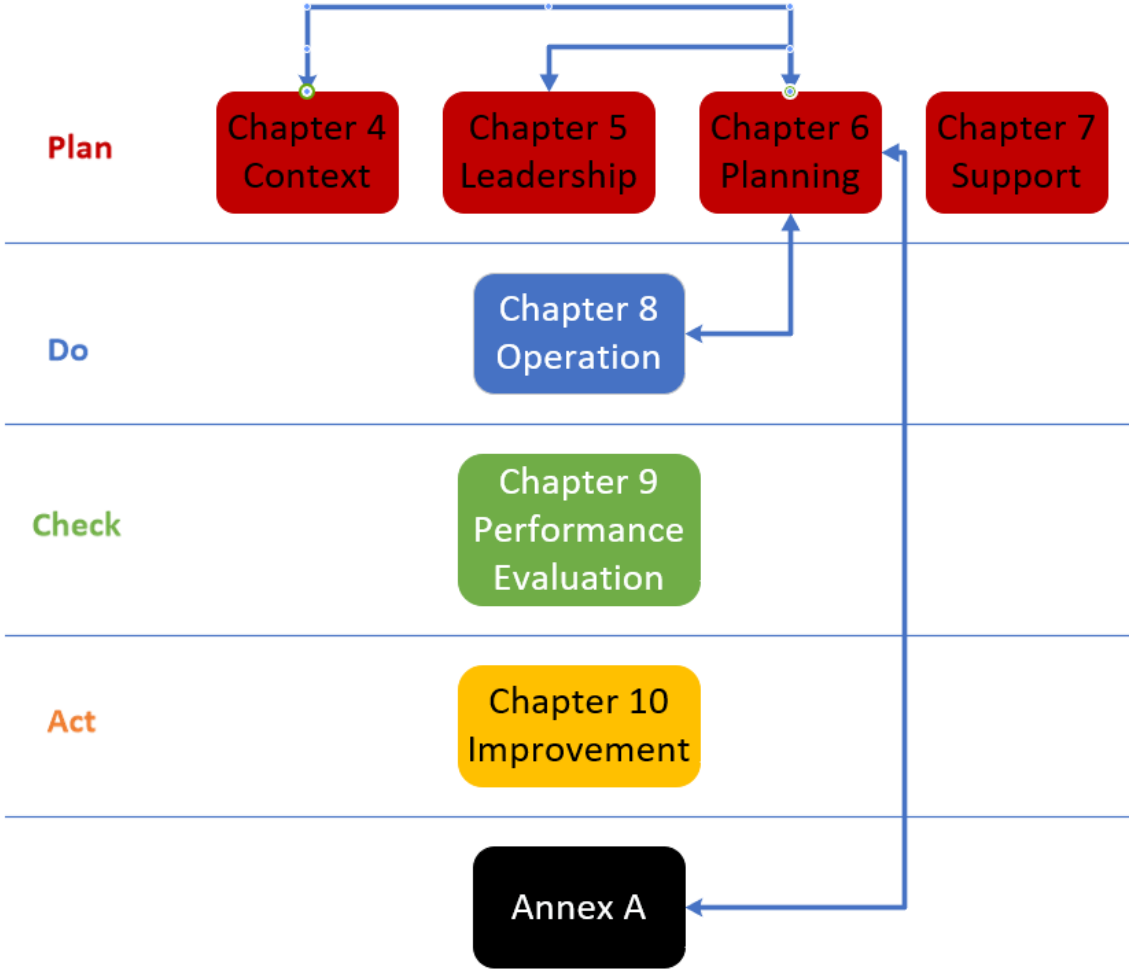
Information Security Management System

- It is an organizational framework linking all the elements relevant to the information security, in order to assure that policies, processes and security objectives are implemented, communicated and assessed.
- It needs to continually improve -> **Deming Cycle**
- It is centered to the risk assessment process -> all decisions are based on the output of this process
- Goal: achieving the optimal CIA **balance**, i.e., ensuring Confidentiality of information, while still ensuring the information remains accessible to authorized persons and is not altered



# The ISO 27001 structure (High Level Structure of All ISOs Standards)

- ISO 27001 composed by
  - 10 Clauses
  - Annex A (normative)
    - Control Objective: statement describing what is to be achieved as a result of implementing controls
    - Control: measure that is modifying risk
  - Applies the PDCA cycle
- Nonconformity: non-fulfilment of a requirement
- Corrective Action: action to eliminate the cause of a nonconformity and to prevent recurrence




# 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

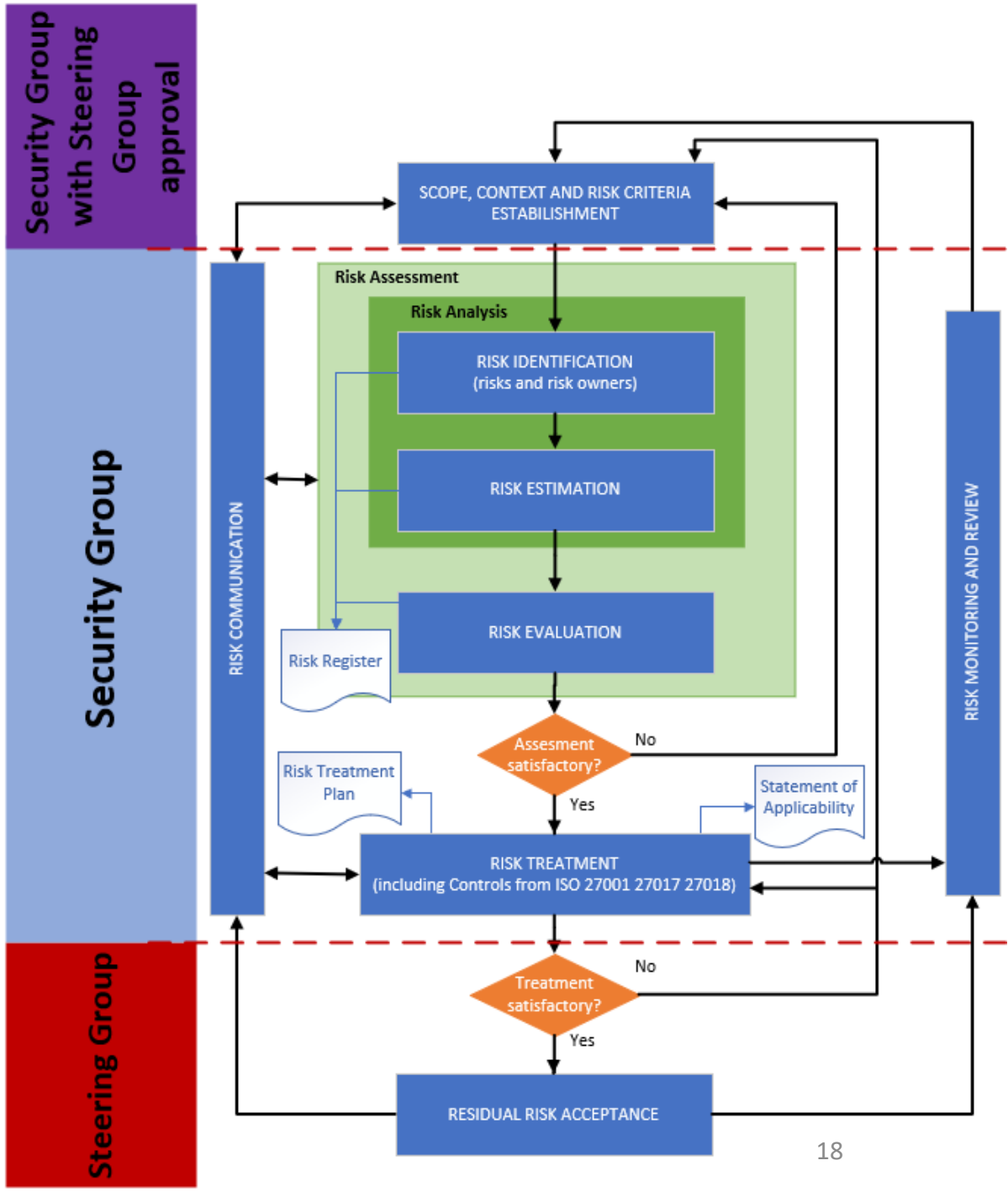
Internal Audit
External Audit

**Act**
(continual improvement)

Effectiveness

# Risk Management Process in EPIC

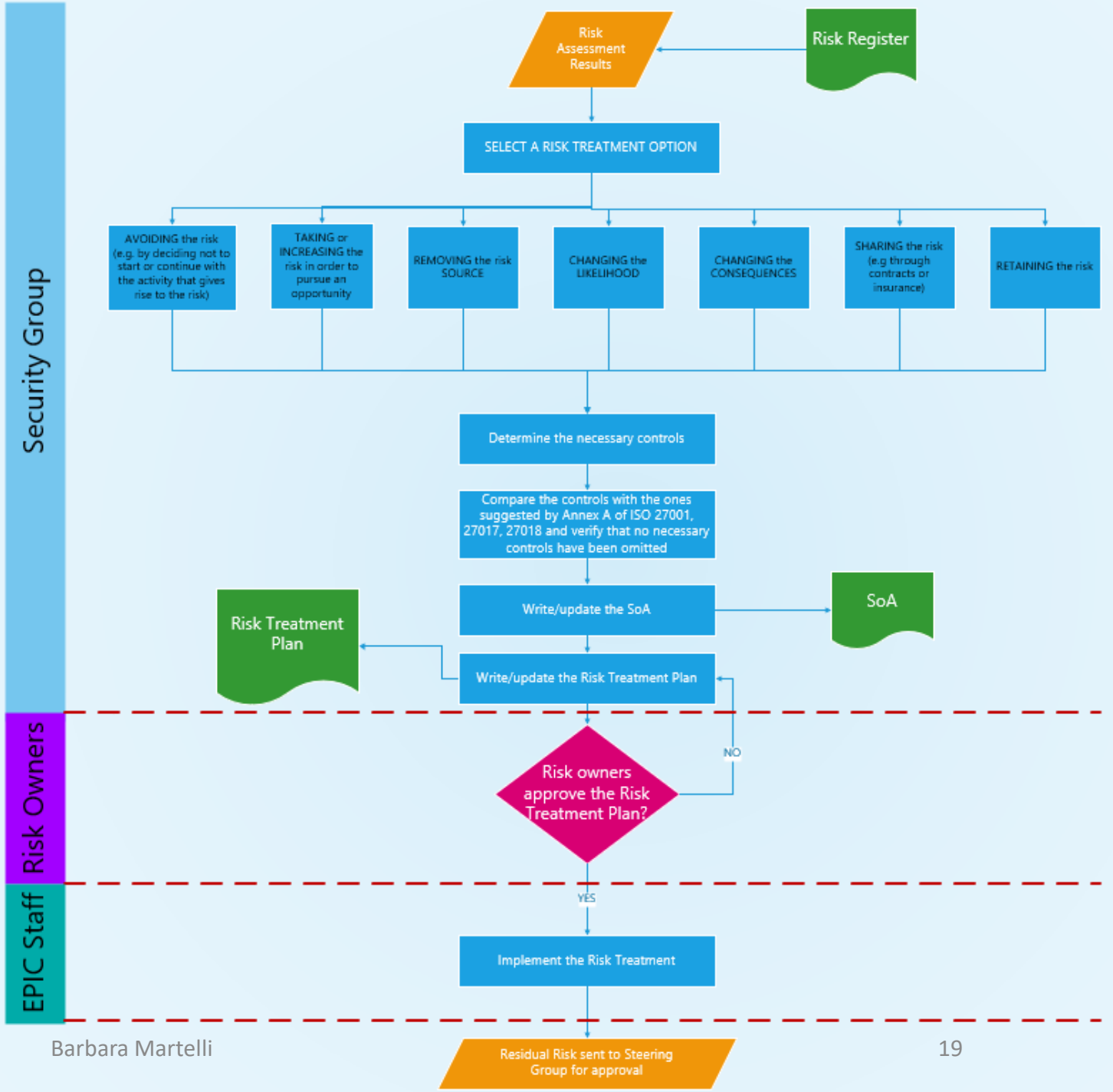
- Inspired by:
  - ISO 27005 guideline with modifications
    - we don't start with asset identification, but we use a scenario-based risk assessment
  - and ISO 31000
- Iterative process aimed at supporting the decision-making process
- In EPIC is performed once a month and whenever a relevant change in the system occurs
  - ISO 27001 clause 8.2 requires to perform it *"at planned intervals or when significant changes are proposed or occur, taking account of the criteria established in 6.1.2 a). "*
- Established criteria to define Risk Owners (person or entity with the accountability and authority to manage a risk), one per Annex A section)



# Information Security Risk Treatment (Clause 6.1.3)

Defined a Risk Treatment process to:

- Select risk treatment options
- Determine the controls that are necessary *(control is something which lowers the probability or the impact of the risk)*
  - *Controls can be freely designed by each organization*
- Compare the controls with those in Annex A and verify that no necessary controls have been omitted
  - Controls objectives in Annex A are not exaustive and additional control objectives may be needed
- Produce a Statement of Applicability (SOA) that contains:
  - The necessary controls
  - Justification for their inclusion
  - Whether the necessary controls are implemented or not
  - The justification for excluding any of the Annex A controls
- Write an information security risk treatment plan
- Obtain risk owners' approval of the information security risk treatment plan and acceptance of the residual IS risks




Do

# EPIC SoA

All Controls from

- ISO 27001 (114 controls)
- ISO 27017 (37 controls)
- ISO 27018 (25 controls)

are defined and applicable

SoA Metrics:

- Level of implementation of each applicable Control (fully, largely, partially, poorly implemented)
- Whole implementation status (sum of Controls implementation levels)

Table 1 Defined controls from ISO/IEC 27001 27017 27018

| ISO/IEC 27001:2013 Control ID | Document | Applicable | Justification |
|---|---|---|---|
| A.05.01.01 Policies for information security | Security Policy | Yes | Risk Assessment Relevant laws, regulations, contracts, agreements |
| A.05.01.02 Review of the policies for information security | Security Policy | Yes | Risk Assessment Relevant laws, regulations, contracts, agreements |
| A.06.01.01 Information security roles and responsibilities | Assegnazione Obiettivi di sicurezza e Controlli People-roles Table | Yes | Risk Assessment Relevant laws, regulations, contracts, agreements |
| A.06.01.02 Segregation of duties | People-roles Table | Yes | Risk Assessment Relevant laws, regulations, contracts, agreements |
| A.06.01.03 Contact with authorities | A06 Internal organization | Yes | Risk Assessment Relevant laws, regulations, contracts, agreements |
| A.06.01.04 Contact with special interest groups | A06 Internal organization | Yes | Risk Assessment Relevant laws, regulations, contracts, agreements |
| A.06.01.05 Information security in project management | A06 Internal organization | Yes | Risk Assessment Relevant laws, regulations, contracts, agreements |
| A.06.02.01 Mobile device policy | A06 Internal organization | Yes | Risk Assessment Relevant laws, regulations, contracts, agreements |
| A.06.02.02 Teleworking | A06 Internal organization | Yes | Risk Assessment Relevant laws, regulations, contracts, agreements |
| A.07.01.01 Screening | A07 Human Resources Security | Yes | Risk Assessment Relevant laws, regulations, contracts, agreements |
| A.07.01.02 Terms and conditions of employment | A07 Human Resources Security | Yes | Risk Assessment Relevant laws, regulations, contracts, agreements |
| A.07.02.01 Management responsibilities | A07 Human Resources Security | Yes | Risk Assessment Relevant laws, regulations, contracts, agreements |
| A.07.02.02 Information security awarness, education and training | Training plan | Yes | Risk Assessment Relevant laws, regulations, |

# Performance Evaluation (Clause 9)

- Requirement: evaluate (and document) the ISMS
  - Performance
  - **Effectiveness**
- Evaluation and monitoring tools
  - Key Performance Indicators (KPI) e.g., systems availability, incident statistics, level of maturity with respect to best practices like ISO 15504 (ITIL (Information Technology Infrastructure Library), COBIT (Control Objectives for Information and related Technology), CMMI (Capability Maturity Model Integration) )
  - Internal audits (at planned intervals, at least once a year)
  - External audits (at planned intervals, at least once a year)
  - Management reviews (at planned intervals, at least once a year)

# Performance Evaluation in EPIC

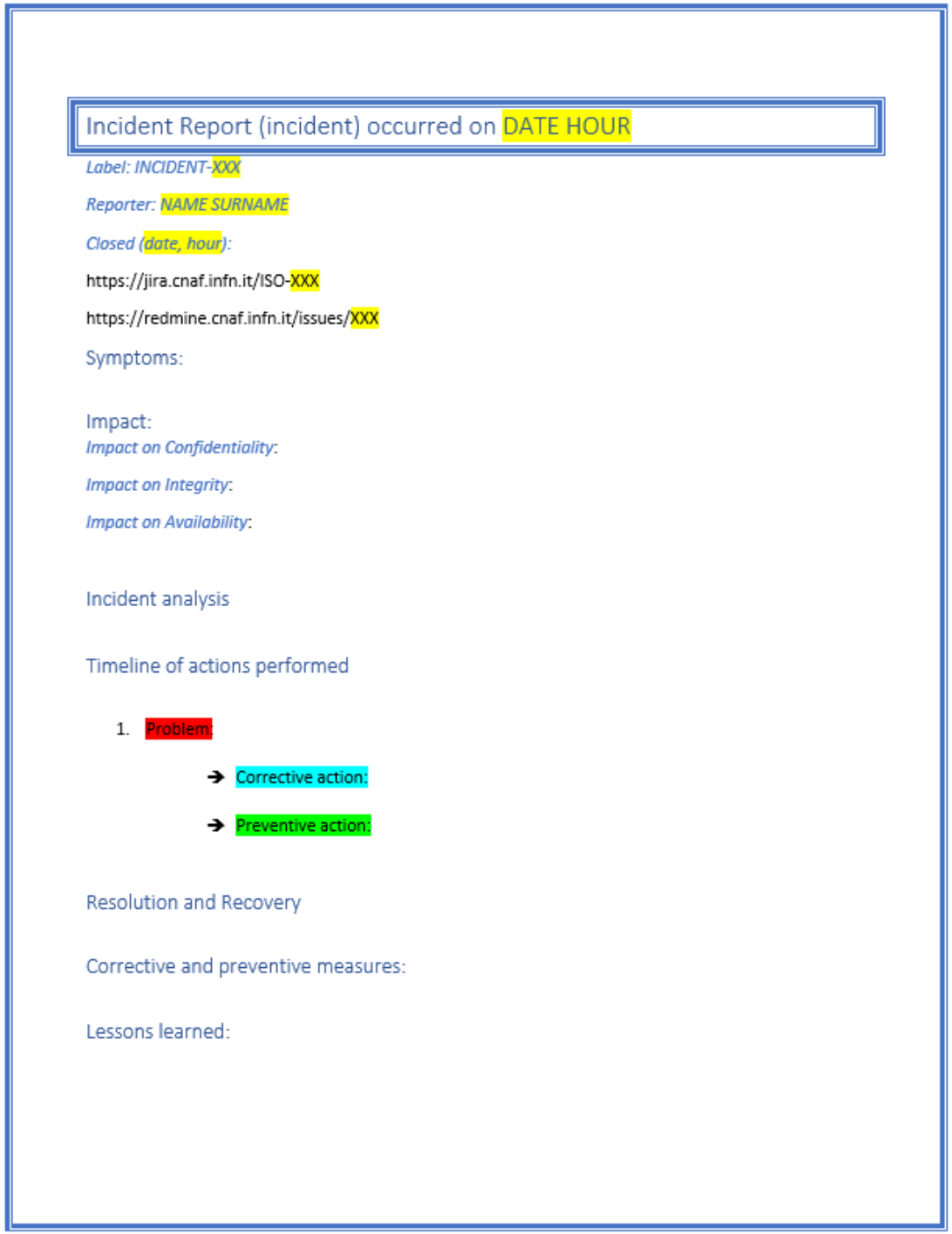
- Management Reviews 4 times a year
  - Identifies improvement areas
  - Review of incident reports, open tasks, nonconformities, corrective actions, monitoring and measurement results
  - Sets Security Objectives and monitors their achievements
- Internal Audit once a year
  - Internal audit plan
  - conducted by an INFN colleague (external to CNAF), with “ISO 27001 lead auditor” certification
  - Identifies and documents nonconformities and opportunities for improvement
- External Audit once a year
  - at present they are conducted by KIWA
  - Identifies and document nonconformities and opportunities for improvement
- Some KPIs:
  - Effectiveness of risk treatment (post-mitigation-risk-level/previous-risk-level)
  - Degree of implementation of SoA controls
  - Number, severity and impact of security incidents
  - Availability of systems

# Improvement (Clause 10)

- When a nonconformity occurs ISO 27001 requires to:
  - Correct it (corrective action)
  - Deal with consequences
  - Evaluate to remove the causes of nonconformity (root-cause analysis)
    - Determine if there are similar nonconformities or if they could occur
  - Implement corrective actions
  - Review the effectiveness of corrective actions
  - Change the ISMS if necessary
- It is required to continually improve the effectiveness of the ISMS

# Improvement in EPIC

- Non-conformities are tracked in Jira and labeled “NC”
  - Corrective actions are recorded in the comments of the Jira tasks
- Information Security Incidents are tracked in Jira and labelled “incident”
  - An incident report is written including impact on CIA  root cause analysis, corrective actions, preventive actions, timeline for recovery, lessons learned
- Incidents and non-conformities are reviewed 4 times a year for improvement



Incident Report (incident) occurred on DATE HOUR

*Label: INCIDENT-XXX*

*Reporter: NAME SURNAME*

*Closed (date, hour):*

https://jira.cnaf.infn.it/ISO-XXX

https://redmine.cnaf.infn.it/issues/XXX

Symptoms:

Impact:

*Impact on Confidentiality:*

*Impact on Integrity:*

*Impact on Availability:*

Incident analysis

Timeline of actions performed

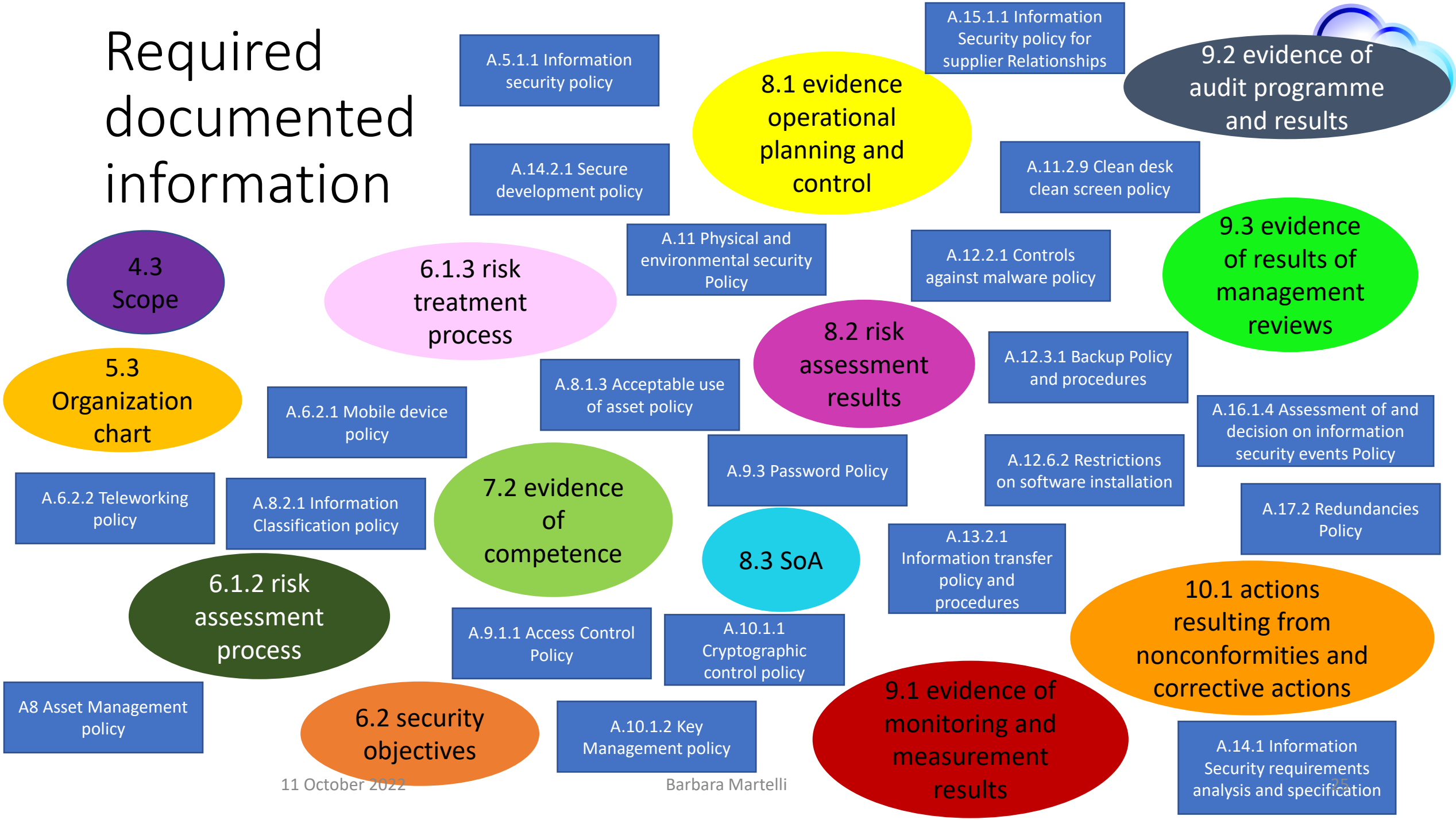
1. Problem:
   - ➔ Corrective action:
   - ➔ Preventive action:

Resolution and Recovery

Corrective and preventive measures:

Lessons learned:

# Required documented information



# Future Work (1/2)
# Expansion of the Management System to the Whole INFN Cloud

- We found that adopting a Management System requires a considerable effort in the first phase, but it gives back a lot of advantages:
  - Clear organizational structure, with clear roles and responsibilities ease the management of complex federations of clouds
    - Policies and risk assessment help identifying priorities and pain points
  - Clear and documented procedures ease the operations
    - Don’t think about word documents, a documented procedure can be a Continuous Integration automated process or an automated script
  - Established risk assessment process helps not to forget any security control
  - Established incident management process helps not repeating the same error twice

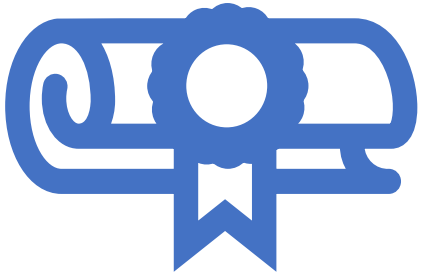
# Future Work (2/2)
# Expansion of the Scope of the ISO Certification

- Adding more sites: Bari is the next one to be certified
- Adding more topics:
  - Quality (ISO 9001)
  - Business Continuity (ISO 22301)
  - Service management (ISO 20000)

Target: obtain the new certifications by october ‘23

# Expansion of the Management System

We decided to extend the Management System to the whole INFN Cloud



This won't be certified, but we expect benefits like the ones obtained in the certified region

Certification is needed to demonstrate to stakeholders the application of the standard, it is possible to adopt the standard without being certified

# Summary and final thoughts

*"Information security is often thought of as being a technical solution. However, the information security that can be achieved through technical means is limited and can be ineffective without being supported by appropriate management and procedures within the context of an ISMS."*

ISO/IEC 27000:2018
https://standards.iso.org/ittf/PubliclyAvailableStandards/index.html

Information Security is not achievable with technical measures alone
you need organizational measures as well

# Summary and final thoughts

- Information Security Management System
  - Process based approach
  - Risk management is the key in the ISMS maintainance
  - Continual iteration of IS processes -> continual improvement
- Requirements are an input of the system, they are not fixed by the ISO standards
  - Availability is key part of information security; an ISMS can be useful even if you don't have confidentiality requirements at all!
    - FAIRness can be another feature guaranteed by an ISMS

# …and certification alone is not enough!


There are plenty of certified ISMS that achieve Information Security only in theory

Focus on effectiveness, not on documentation