EOSC4Cancer
A European-wide foundation to accelerate Data-driven Cancer Research

This project has received funding from the European Union’s Horizon Europe Programme under GA 101058427 — EOSC4Cancer — HORIZON-INFRA-2021-EOSC-01
Context

European Health Data Space
European Cancer Mission

Alignment

Open data standards

Stakeholders

...
EOSC4Cancer
A European-wide foundation to accelerate Data-driven Cancer Research

➔ EOSC4Cancer as provider of the infrastructure for the exploitation of cancer data for the EU Cancer Mission.

➔ EOSC4Cancer brings together comprehensive cancer centres, research infrastructures, leading research groups, and major computational infrastructures across Europe to make the exploitation of the data possible.

➔ EOSC4Cancer will prepare EOSC services for cancer research and enrich EOSC with data, tools and services from the cancer community.

➔ **Start:** 1 September 2022 (30 months)

➔ **Consortium:** 28 full beneficiaries, 1 associated partner & 5 affiliated entities.
Partners, associates, and affiliated entities
Objectives

➔ Enable **storage, access, sharing, analysis** and processing of research data and other digital **research objects** from basic and clinical cancer research.

➔ Mobilise, **interconnect** and **interoperate datasets** relevant in cancer research.

➔ Make cancer research **data** and **analysis systems** accessible to basic and clinical scientists in the most used cancer **analysis portals**.

➔ Integrate digital tools, data analytics and Artificial Intelligence/Machine Learning tools for the analysis of cancer data in the cancer analysis portals.

➔ Contribute to the **European Health Data Space** (EHDS), the Horizon Europe **European Open Science Cloud** (EOSC) Partnership and the **Cancer Mission**
Expected Outcomes and Impact

➔ **Facilitate Cancer Research across Member States and Associated Countries**

**Result #1:** A platform that will enable *storage, sharing, access, analysis* and processing of research data and other digital research objects from *basic and clinical cancer research*.

**Result #2:** Mobilisation, interconnection and interoperation of *datasets relevant in cancer research*. Contribute with protocols and operating procedures to facilitate the progressive adoption of the *FAIR principles* across data sources but also for research software.

➔ **Researchers, healthcare professionals, cancer patients, and survivors contributing to cancer research**

**Result #3:** Cancer research data and analysis systems made *easily accessible* to basic and clinical scientists in open cancer analysis portals.
Expected Outcomes and Impact


➔ **Contribute to the Horizon Europe EOSC Partnership and other relevant partnerships related to cancer research.**

Result #5: Roadmap for EOSC support to an European Cancer Data Space in the European Health Data Space (EHDS) in partnership with the Cancer Mission. Participation of patient/survivors associations in the project’s Stakeholder forum and in the Scientific Advisory Board together with the Ethics Advisory Board.
The patient journey

Prevention
Cancer origin
- Cancer registries
- Environment (pollutants)
- Social data
- Geolocalisation

Diagnostics
Primary tumours
- Screening programmes
  - Genomics data
  - Medical imaging
  - Medical data (EHR structured data)

Treatment
Metastatic Cancers
- Patient level:
  - Genomics data, Imaging (digital pathology), and EHR data (structured data)
  - Other research data: liquid biopsy, animal models, drug screening
  - Non patient specific data: animal models, drug screening, etc

Cancer Research
- Cancer clinical trials (descriptors - metadata)
- Patient level genomics and medical data
- Actionable data (regulated): markers, drugs, and treatments

Research Software for epidemiological level analysis
Software, Workflows and Portals for patient level clinical research
Clinical Decision Support Systems
Technical overview

Databases and repositories
- Biobanks
- EHRs
- Federated EGA
- Imaging Data
- Cancer Registries

Knowledge databases
- Approved Markers
- Ongoing Clinical Trials
- Actionable Genes

User requests flow

WP1
- Research Portals & Clinical Decision Systems
  - MTB
  - cBioPortal
  - Galaxy
  - DataShield
  - OpenVRE

WP2
- Combined data for individual patient/s

WP3
- Based on software containers and reference analytical workflows

WP4
- Use Cases

WP5
- Training

WP6
- Community Engagement

Roadmap

Data flow
Technical WPs overview

- **WP1** Data spaces
  - T1.1 Cohort descriptors
  - T1.2 Synthetic data
  - T1.3 Interoperable access
  - T1.4 Federated data flow

- **WP2** Harmonization and Interoperability
  - T2.1 Clinical metadata mining, archiving and interoperability
  - T2.2 Multi-omics and imaging
  - T2.3 Discoverability

- **WP3** Federated analysis and visualization
  - T3.1 Clinical trial databases and CDSS
  - T3.2 Data portals
  - T3.3 Virtual research environments

- **WP4** Use cases
  - Identify data resources
  - Identify analysis tools and visualization
  - Data access mechanisms

- **Metadata descriptors**
Use-cases overview

T4.1 (Lead: CNR IEOS) Cancer risk identification and prevention by linking environmental data to cancer registry data

T4.2 (Lead: NKI) Data driven optimisation of cancer screening programs

T4.3 (Lead: UP) Data-driven treatment selection for localised tumours with multiple patient-derived data types

T4.4 (Lead: NKI) Data-driven treatment selection for localised tumour: improving the treatment of colorectal cancer by the inclusion of circulating DNA information

T4.5 (Lead: VHIO) Connecting omics data from multiple sources to a Clinical Decision Support System (CDSS) for precision treatment of metastatic CRC
Kick-off meeting Outcomes: Squads

Squad #1

T4.1 (Lead: CNR IEOS) Cancer risk identification and prevention by linking environmental data to cancer registry data

T4.2 (Lead: NKI) Data driven optimisation of cancer screening programs

Squad #2

T4.3 (Lead: UP) Data-driven treatment selection for localised tumours with multiple patient-derived data types

T4.4 (Lead: NKI) Data-driven treatment selection for localised tumour: improving the treatment of colorectal cancer by the inclusion of circulating DNA information

Squad #3

T4.5 (Lead: VHIO) Connecting omics data from multiple sources to a Clinical Decision Support System (CDSS) for precision treatment of metastatic CRC

WP1
Data spaces

WP2
Harmonization and Interoperability

WP3
Federated analysis and visualization
Collaborations with other INFRAEOSC projects

➔ **BY-COVID** and **EOSC4Cancer** share the same conceptual architecture. Thus, lesson learnt in BY-COVID, e.g. on data access and mobilization, can be leveraged by EOSC4Cancer. Similarly, any technological development can be taken up, maintained and extended by EOSC4Cancer.

➔ **EuroScienceGateway** will leverage existing technologies (e.g. Galaxy, Pulsar, Dirac) and extend them to allow smart workflow execution scheduling around Europe. **EOSC4Cancer** represents the ideal scenario to tackle the challenges on analyzing sensitive data across different places.
Top 3 technical challenges

- Accessibility of cancer-related datasets
- Harmonization and interoperability of cancer-data resources and analysis systems (FAIR principles)
- Integration of digital tools, data analytics, and Artificial Intelligence/Machine learning tools in the cancer analysis portals and federated systems

Top 3 engagement challenges

- Align and engage with large international and EU coalitions, e.g., ICGC-Argo, GA4GH, 1+MG/B1MG, UNCAN.eu, canSERV
- Contribute to EOSC services, the European Health Data Space (EHDS) and the Cancer Mission
- Learn from and create value for five use cases following the patient journey, laying the foundation for prospective studies
Thank you!

Contact: info@eosc4cancer.eu

Follow us in social media

This project has received funding from the European Union’s Horizon Europe Programme under GA 101058427 — EOSC4Cancer — HORIZON-INFRA-2021-EOSC-01