

Validating AI models for prostate cancer screening across institutions (MCVAL)



BBMRI-ERIC®

biobanking for a healthier world

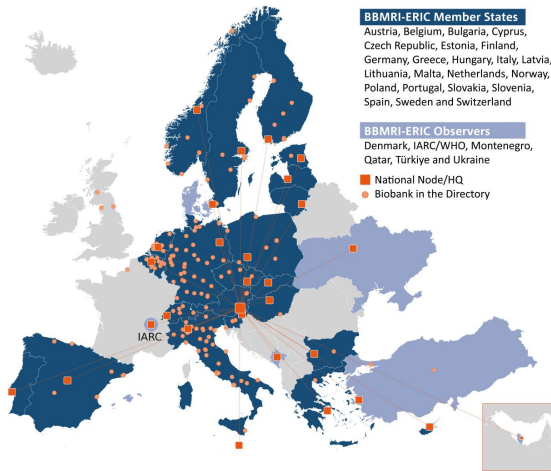
Assoc. Prof. Petr Holub, Ph.D.
Chief Information Officer (CIO)
BBMRI-ERIC EOSC Node Coordinator

BBMRI IS A DISTRIBUTED ERIC

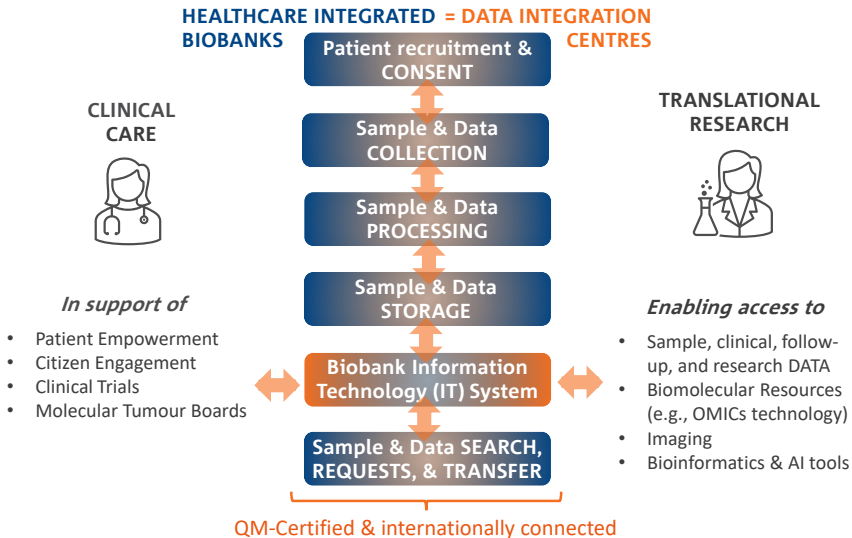
Operating in health and life sciences

BBMRI-ERIC

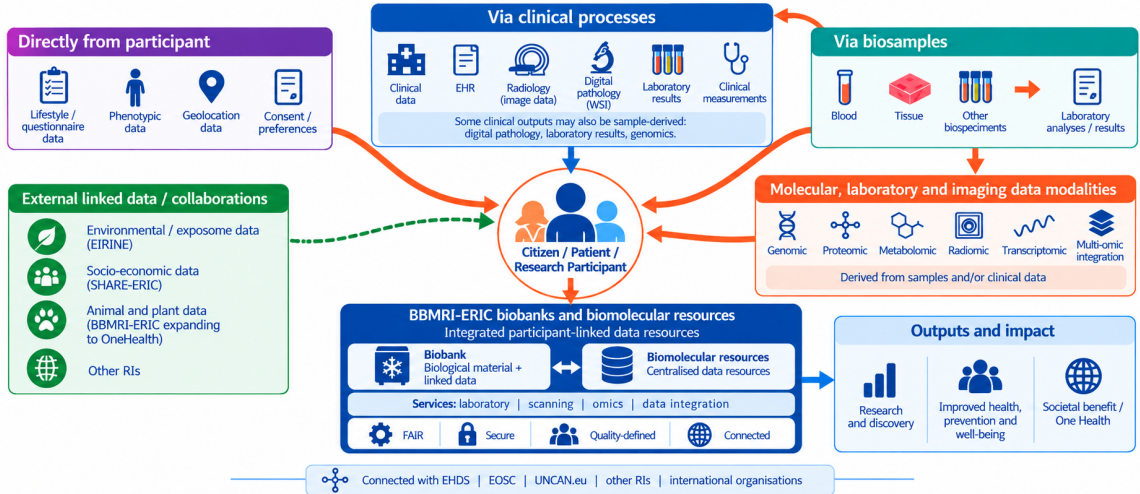
- 28 Member and Observer countries and IARC/WHO
- Facilitates access to **biobanks** and **biomolecular resources**, advancing health, life science research and innovation
- Brings together a **community** of > 500 biobanks; 3 Biotechnology Expert Centres; > 30 ELSI experts; > 130 IT experts and > 300 Quality Management experts
- **Core operations and services:** Scientific Affairs (Biobanking Development, IT, Quality Management, ELSI) and Central Services & Functions (Public Affairs, Outreach, Education & Communication, Finance & Project Management)



THE LOCAL BIOBANK ECOSYSTEM



Data Modalities in BBMRI-ERIC



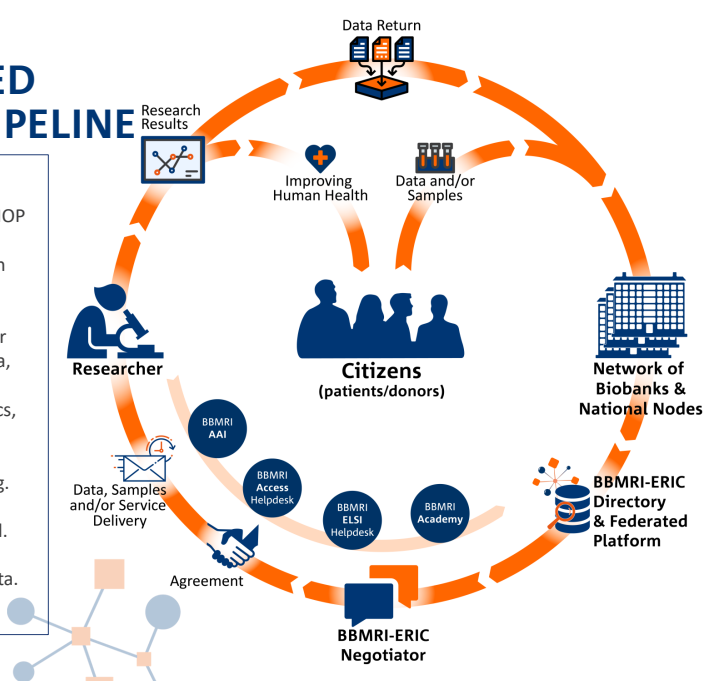
BBMRI'S FEDERATED ACCESS PIPELINE

BBMRI-ERIC's MIABIS*

FAIR Data include:

- >8k harmonised OMOP variables of clinical phenotype data with biochemical measurements,
- >10PB whole slide or radiology image data,
- OMICs data (> 155K genomes, proteomics, metabolomics etc.),
- National Registries connections with e.g. survival data and
- other data types incl. surveys, genealogy, patient-provided data.

*MIABIS: Minimum Information
ABout Biobanking data Sharing



Training & Education in:

- ELSI
- IT
- QM
- Biobank Development
- Health & Life Science Research/Development
- Outreach/Engagement

BBMRI-ERIC Directory is largest biobank catalogue worldwide:

- ~500 biobanks
- 3,768 collections
- 33 countries.

BBMRI-ERIC Federated Platform features:

- 460,000 donors with
- 155,000 genomes across
- 10 EU countries today.

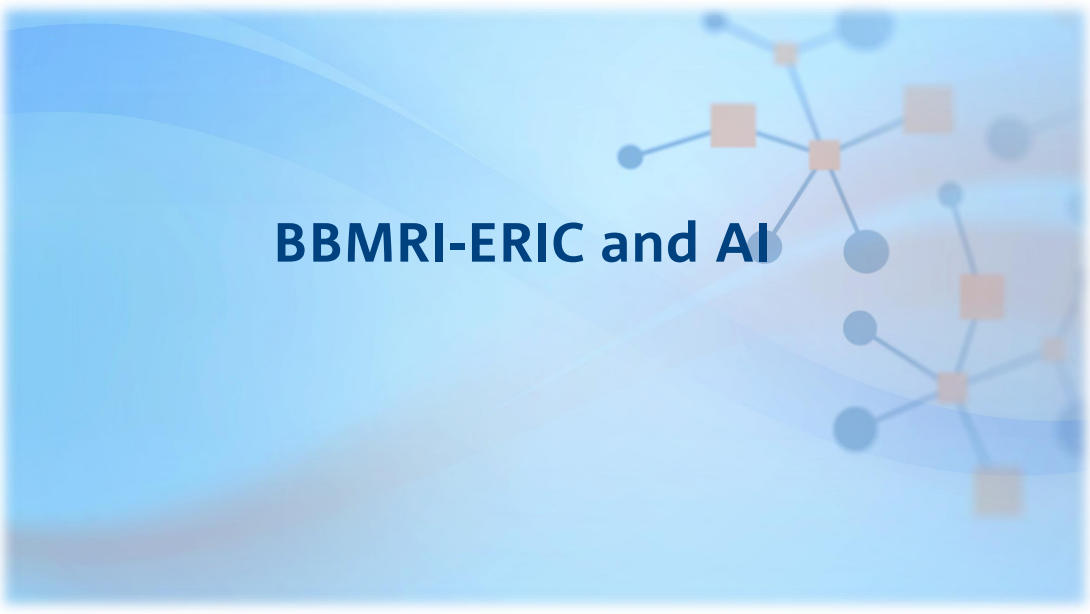
eOSC Node | BBMRI-ERIC

Health & Food

- ▶ **One of 13 first-wave EOSC Federation nodes** (besides EU Node) since 2025
- ▶ **One of initial 14 EOSC Federation nodes in production** since October 8, 2026
- ▶ **Focus on providing access and processing sensitive/health data**, preferably via Secure Processing Environments
- ▶ **Collaboration with other nodes on service chains:**
 - **infrastructures providing sensitive data processing capacities:** EOSC Italy (ICSC/INFN), EOSC Czechia (SensitiveCloud), EOSC Netherlands (SURF), ...
 - **other nodes providing complementary data** – e.g., ENVRI or national nodes such as EOSC Poland or EOSC Slovakia

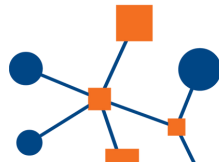


BBMRI-ERIC and AI

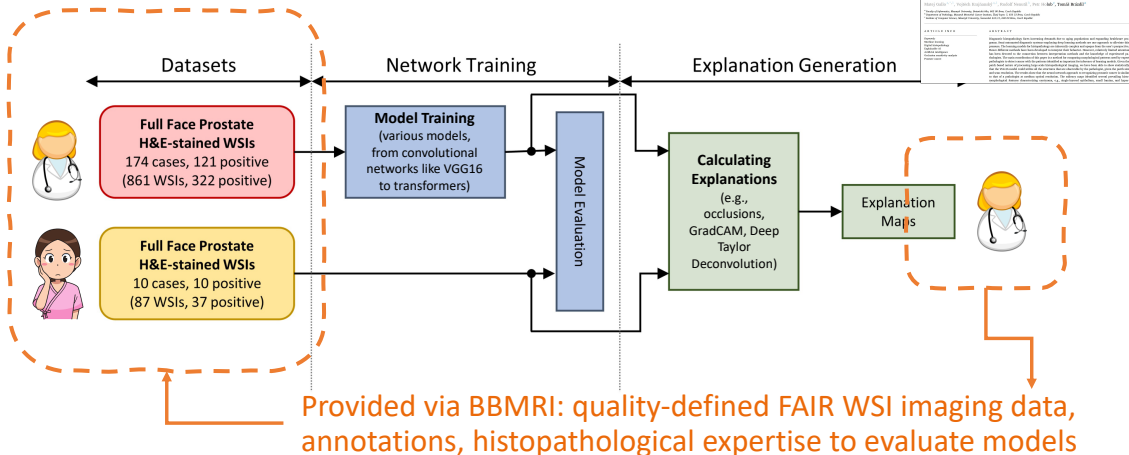


BBMRI-ERIC & AI

- ▶ **BBMRI-ERIC as a data provider for clinically-relevant AI**
 - sample-derived data (imaging, omics, ...) and associated data (clinical, phenotype, exposure, ...) for **model training**
 - environment to **validate AI models** on unseen data – **single-centric** or **multi-centric**
- ▶ **BBMRI-ERIC as an expertise provider for clinically-relevant AI**
 - identification of clinical experts
- ▶ **BBMRI-ERIC as a host of AI groups – to improve services of the RI, e.g.,**
 - **quality checking**
 - **providing descriptors** for data sets **to facilitate discovery**
 - **generate synthetic data** to speed up or proxy access

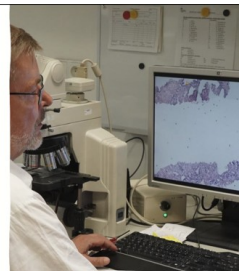
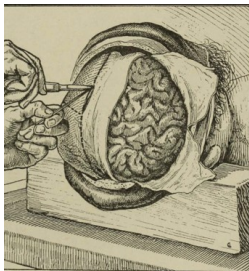


How is BBMRI-ERIC used for eXplainable AI?



MCVAL EOSC Federation Use Case

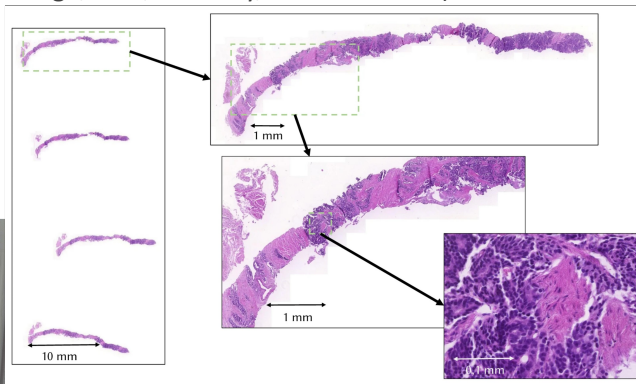
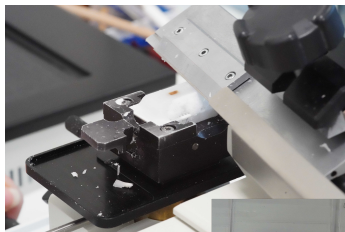
- ▶ **Evolution of pathology:** from examining dead bodies to **critical cancer diagnostics**
- ▶ Projected digital pathology growth: \$1.2B (2024) to \$2.6B (2032) (Delvens [↗](#))



MCVAL EOSC Federation Use Case

Example of prostate slide

- Statistically: about 1 in 8 men will be diagnosed with prostate cancer during their lifetime (risk varies with age, race/ethnicity, other factors...).



- 

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



bioRxiv

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

bioRxiv

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



EOSC Multi-Centric AI Validation Demonstrator

Roles – 36 experts involved!

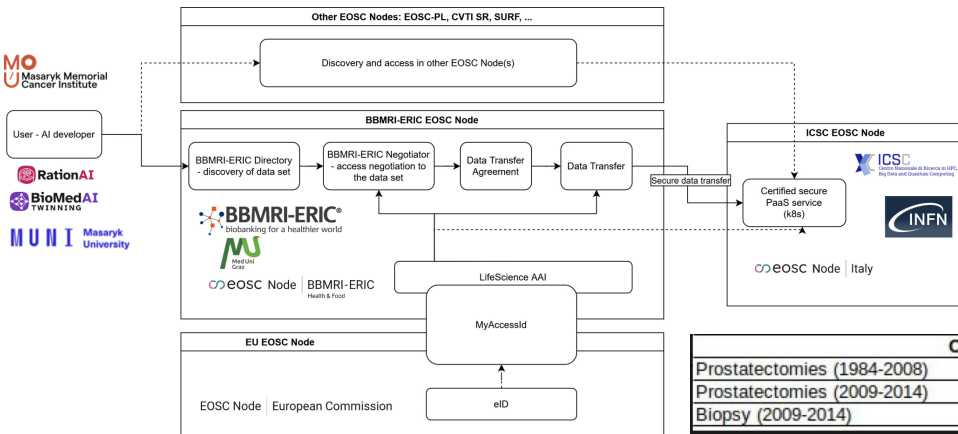
- AI researcher
- Secure compute node
- Sensitive data provider
- AI validation service provider

Participants

- MMCI/MU
- ICSC/INFN
- MUG
- BBMRI-ERIC

Additional EOSC Federation Sensitive Data WG Participants

- EOSC-SK
- EOSC-PL
- EOSC-NL (SURF)
- EOSC-FI



	Cases	Slides
Prostatectomies (1984-2008)	5525	104335
Prostatectomies (2009-2014)	794	28.786
Biopsy (2009-2014)	331	11.193

Lessons Learned

- ▶ **EOSC helps researchers to orchestrate complex workflows spanning different EOSC nodes**
- ▶ **Sensitive data constitutes only a subset of EOSC**
 - collaboration stimulates **interoperability** and **fitness for purpose**: authentication/authorization infrastructures, secure processing environments, secure data transfers
 - implementing secure data transfer mechanisms between data sources and adequately sized and secured infrastructures
 - "computation to data" works only for some scenarios
- ▶ **MCVAL demonstrator have showed how EOSC can be used to bring together medical data sources in BBMRI-ERIC with eInfrastructure/HPC capacit providers**



THANK YOU!



BBMRI-ERIC®
biobanking for a healthier world

petr.holub@bbmri-eric.eu