BioMedAI TWINNING

RationAI

BioMedAI · RationAI

Clinically Focused AI

Ostrava, 2024-12-21

doc. Petr Holub, Ph.D., doc. Tomáš Brázdil, Ph.D.

Institute of Comptuter Science · MUNI

MUNI ICS

Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Funded by the European Union

RationAI – BioMedAI

- **RationAI**
  - laboratory focusing on rational application of AI with focus on explainability and trustworthiness
  - founded in 2019 as a shared laboratory between FI and ICS
    - FI focuses on core research
    - ICS focuses on development of infrastructure for AI, provides project management and back-office

BioMedAI TWINNING

RationAI

MUNI ICS

Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Funded by the European Union

RationAI – BioMedAI

- **BioMedAI**
  - center bringing together RationAI with a image analysis, visualization, and NLP laboratory for **eXplainable AI (XAI) research in biomedical applications** with specific focus on **digital pathology**
  - partnering with Masaryk Memorial Cancer Institute, Medical University Graz, and Technical University Berlin
  - founded in 2023 as a consequence of BioMedAI Twinning project

BioMedAI TWINNING

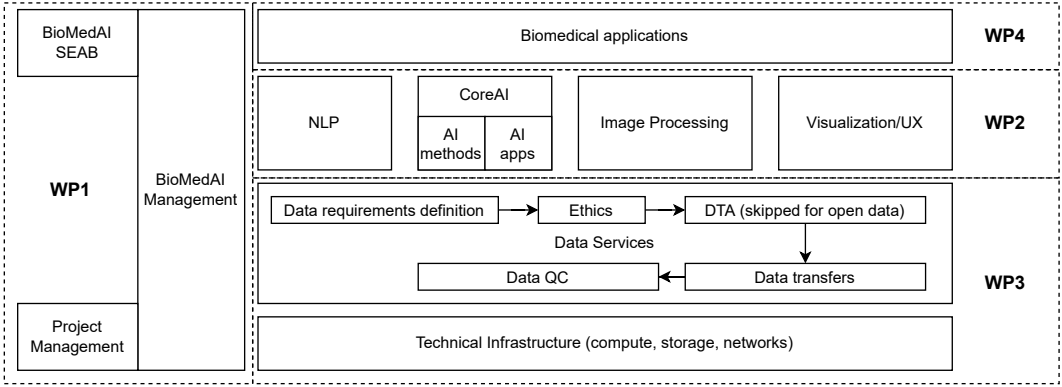
RationAI

Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

MUNI ICS

Funded by the European Union

BioMedAI Structure



BioMedAI TWINNING

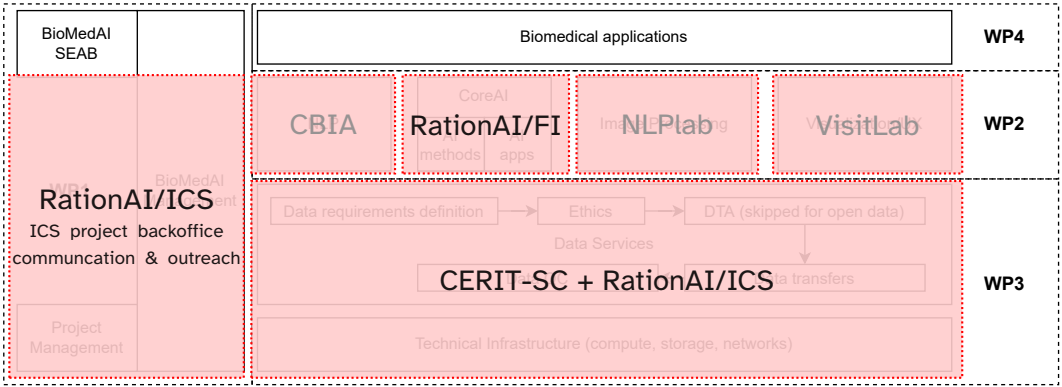
RationAI

MUNI ICS

Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Funded by the European Union

BioMedAI Structure



BioMedAI TWINNING
RationAI
MUNI ICS
Funded by the European Union

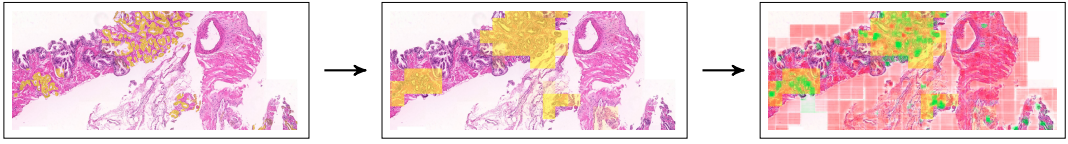
XAI Research Domains - I

- **Digital pathology – primary focus**
  - XAI analysis of large-scale imaging
  - screening and diagnostic applications
  - application of explainability to develop trust by medical professionals and to understand boundary conditions of applicability

https://doi.org/10.1016/j.nbt.2023.09.008
ELSEVIER
New Biotechnology
Volume 78, 25 December 2023, Pages 52-67

Shedding light on the black box of a neural network used to detect prostate cancer in whole slide images by occlusion-based explainability

Matej Gallo [a 1], Vojtěch Krajňanský [a 1], Rudolf Nenutil [b], Petr Holub [c], Tomáš Brázdil [a]



BioMedAI TWINNING

RationAI

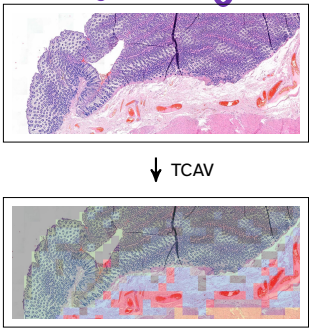
MUNI ICS

Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Funded by the European Union

XAI Research Domains - II

- **Focus topics**
  - **fast and efficient training** of AI models to different problems in digital pathology – **transfer learning** and **domain adaptation**
  - **active learning** mechanisms – including novel visualization and HCI methods
  - **post-hoc model-independent explainability** methods
  - **extracting "knowledge"** from models (e.g., Testing with Concept Activation Vectors – TCAV)
  - **anomaly detection** models – quality control, data synthesis
  - feature extraction of **efficient discovery** of data in large data sets



BioMedAI TWINNING
RationAI
MUNI ICS
Funded by the European Union

Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

XAI Research Domains - III

- **collaborations:**
  - Masaryk Memorial Cancer Institute: Dept. of Pathology
  - Medical University Graz: Human Centered AI Group, Zatloukal Group
  - Charité Berlin + EMPAIA International
  - Technical University Berlin: Distributed Artificial Intelligence Laboratory (DAI Labor)
  - Faculty Hospital Brno: Inst. of Pathology
  - Institute for Clinical and Experimental Medicine (IKEM)
  - BBMRI-ERIC: European Research Infrastructure on Biobanking and Biomolecular Resources

BioMedAI TWINNING
RationAI
MUNI ICS
Funded by the European Union
Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

XAI Research Domains - IV

- **Obesitology & laboratory medicine**
  - analysis and predictions of time series data
  - **collaborations:** (i) 1st Medical Faculty at Charles University, and (ii) Institute of Laboratory Medicine at Faculty Hospital Brno
- **Public health**
  - data-quality analysis of national medical registers
  - **collaborations:** ÚZIS (Institute of Health Information and Statistics of CZ)

BioMedAI TWINNING
RationAI
MUNI ICS
Funded by the European Union

Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Infrastructure R&D Domains - I

- **Co-initiated development of SensitiveCloud**
  - ISO 27001 certified environment for hosting and processing sensitive data
  - Kubernetes-based processing environments, piloting feasibility using RationAI group
- **Development of Kubernetes-based RatioAI pipeline and RatioViz visualization**
  - support for interactive development and batch HPC mode
  - setup of reproducible AI pipeline management based on MLflow and visualization
  - close collaboration between pipeline/visualization developers and infrastructure developers – architectural co-design

BioMedAI TWINNING

RationAI

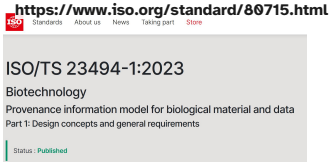
MUNI ICS

Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Funded by the European Union

Infrastructure R&D Domains - H

- **Provenance modeling and standardization**
  - collaboration with BBMRI-ERIC on standardization of provenance in ISO TC/276 (Biotechnology) WG5 (Data Integration) – ISO 23494 Series
  - provenance model for sensitive data in distributed environments
  - application of provenance to the whole digital pathology chain: patient → sample → histopathological slide → whole-slide image → training of AI model → testing of AI model → clinical validation



BioMedAI TWINNING

RationAI

MUNICS

Infrastructure R&D Domains - HI

- **Privacy risks of privacy sharing**
  - analysis of risks of digital pathology imaging sharing
  - development of guidelines for anonymization of digital pathology imaging

- **Optimizations of imaging formats**
  - collaboration with Comprimato spin-off on using JPEG 2000 compression for substantial reduction of data volumes in digital pathology
  - development of DICOM support

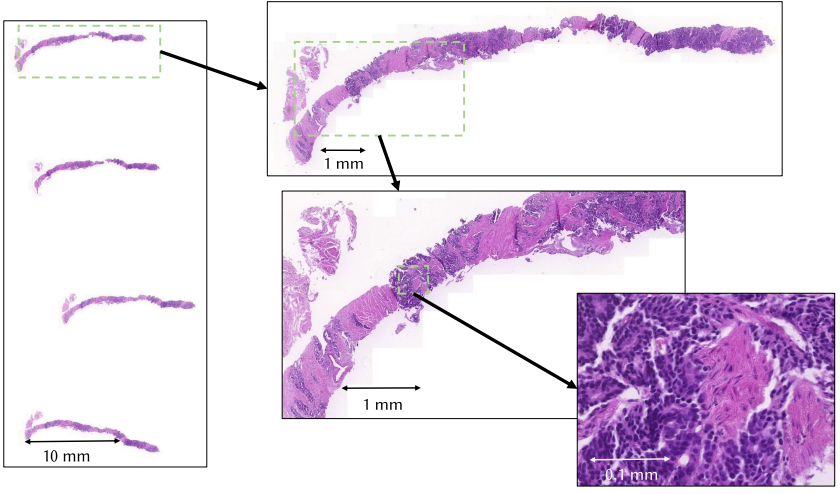
BioMedAI TWINNING

RationAI

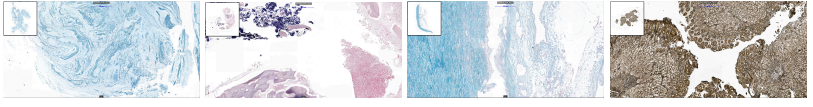
MUNI ICS

Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Funded by the European Union



(a)



(b)

Data Management - I

- **Setup of contracting with new data/research partners**
  - ongoing expansion of collaborating hospitals: national and international
- **Development of FAIR/FAIR-Health data management plans**
  - FAIR: findability, accessibility, interoperability, reusability
  - FAIR-Health: quality (starting from traceability :)), (effective) privacy protection, incentives

Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Funded by the European Union

Data Management - II

- **Routine data management**
  - contracting access to data in large organizations
  - implementation of secure compartmentalization of storage
  - optimization of storage performance
    (TB–PB range of datasets)
  - infrastructural monitoring of compliance
  - operational data management done within RationAI group

BioMedAI TWINNING
RationAI
Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.
MUNI ICS
Funded by the European Union

Data Management - III

- **Data pipeline:**
  - tiling → labeling → loading → training ...all captured in mlFlow
  - loading and processing in SensitiveCloud storage
- **Examples of data sets**
  - 123 TB biobank data set (MMCI/BBMRI)
  - 3 TB prostate data set (MMCI/BBMRI)
  - 1.5 PB colorectal cancer data set (MUG/BBMRI)

BioMedAI TWINNING RationAI

Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

MUNI ICS

Funded by the European Union

Output Focus

- **Applied research results**
  - RationPath pathology toolset
    - efficient in adapting to different problems
  - RatioViz
    - xOpat client visualization with different compatible servers, “case viewer’, annotation tooling
    - integrated in BBMRI-ERIC toolset with cBioPortal
  - software developed as open source (https://github.com/RationAI/ once published)
- **Trained models in internal clinical validation at MMCI**

BioMedAI TWINNING

RationAI

MUNI ICS

Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Funded by the European Union

International Leadership

- **EOSC**
  - PH: lead of Health Data Task Force, writer of EOSC Handbook, WP6 lead of EOSC-Life (FAIR data services)
- **European Health Data Spaces**
  - PH: consultant to EHDS architect since 2020, leader of WP8 of EHDS2Pilot (data quality, privacy/security, and processing), contributor to TEHDAS2 and QUANTUM
- **ISO**
  - PH: initiator and project lead of ISO 23494, contributor to ISO 20691

BioMedAI TWINNING

RationAI

MUNI ICS

Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Funded by the European Union

INTEREST IN OPEN SCIENCE II

Topics - I

- **Traceability and quality of health-related data**
  - provenance of data sets back to source
  - provenance of AI models
  - including support for sensitive medical data (provenance not leaving source organizations)
    - provenance not leaving source organizations
  - quality checking (piloted on clinical and imaging data)
  - compliant with ISO 23494

BioMedAI TWINNING

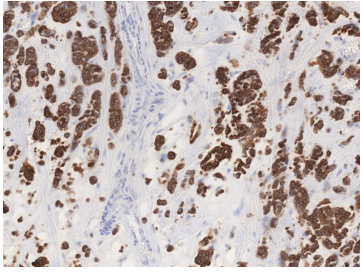
RationAI

Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

MUNI ICS

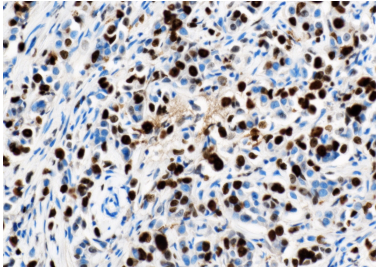
Funded by the European Union

Topics - II



epithelium

×



Ki67-positive nuclei

BioMedAI TWINNING

RationAI

Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

MUNI ICS

Funded by the European Union

Topics - III

- **Alignment to development of European Health Data Spaces**
  - working on behalf of Czech Ministry of Health in TEHDAS2
  - privacy-preserving data sharing (anonymization/pseudonymization)
  - discovery and access management
  - provisioning of Secure Processing Environments (SPEs)

BioMedAI TWINNING

RationAI

MUNI ICS

Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Funded by the European Union

Topics - IV

- **Licensing of AI models**
  - what is legaly ’an AI model’?
  - open licensing mechanisms
  - contract clauses with employees and students contributing to AI models

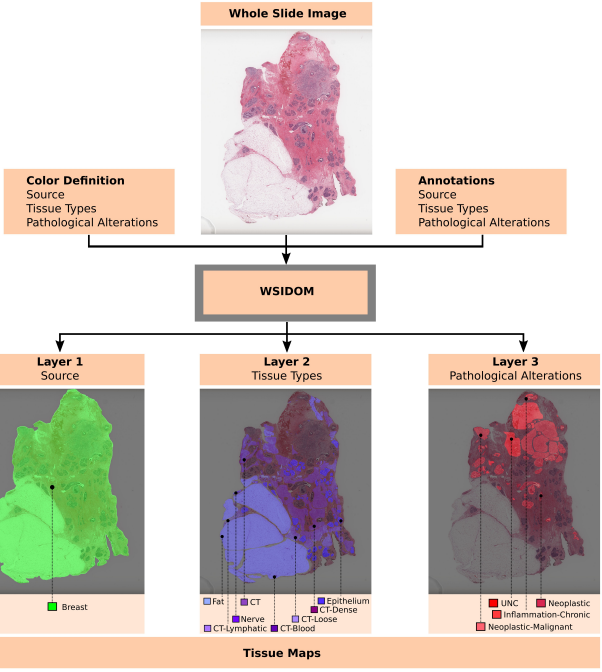
BioMedAI TWINNING
RationAI
MUNI ICS
Funded by the European Union
Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Topics - V

- **Repositories for AI-ready data with controlled access**
  - metadata models
    - information extracted from data
    - information extracted from provenance
  - interfacing repositories to SPEs (e.g., Kubernetes-bases in CERIT-SC)
  - focusing on ’our domains’ – clinical data ((un)structured, imaging, clinical, omics)

BioMedAI TWINNING
RationAI
MUNI ICS
Funded by the European Union
Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Topics -



BioMedAI TWINNING

Views and opinions expressed are those
European Research Executive Agency (RE

Funded by the European Union

Topics - VII

- **Repositories for AI models**
- **Effective storage of large-volume data in specific domains**
  - histopathological and other large-scale imaging data
  - time series and clinical data

BioMedAI TWINNING
RationAI
MUNI ICS
Funded by the European Union