Privacy-Preserving Analytics in Medicine (PrivateAIM)

Prof. Dr. Fabian Prasser
Prof. Dr. Oliver Kohlbacher
Prof. Dr. Daniel Rückert

GEFÖRDERT VOM
Bundesministerium
für Bildung
und Forschung

Project Overview

- Module 3 – Method Platform within the coming phase of the MII
- Goal of the project:

  „The goal of PrivateAIM is to develop a federated machine learning (ML) and data analytics platform for the Medical Informatics Initiative (MII), where analyses come to the data instead of data coming to the analyses."

- "Code to data" paradigm – data remains where it is to reduce potential privacy impact and resolve legal issues

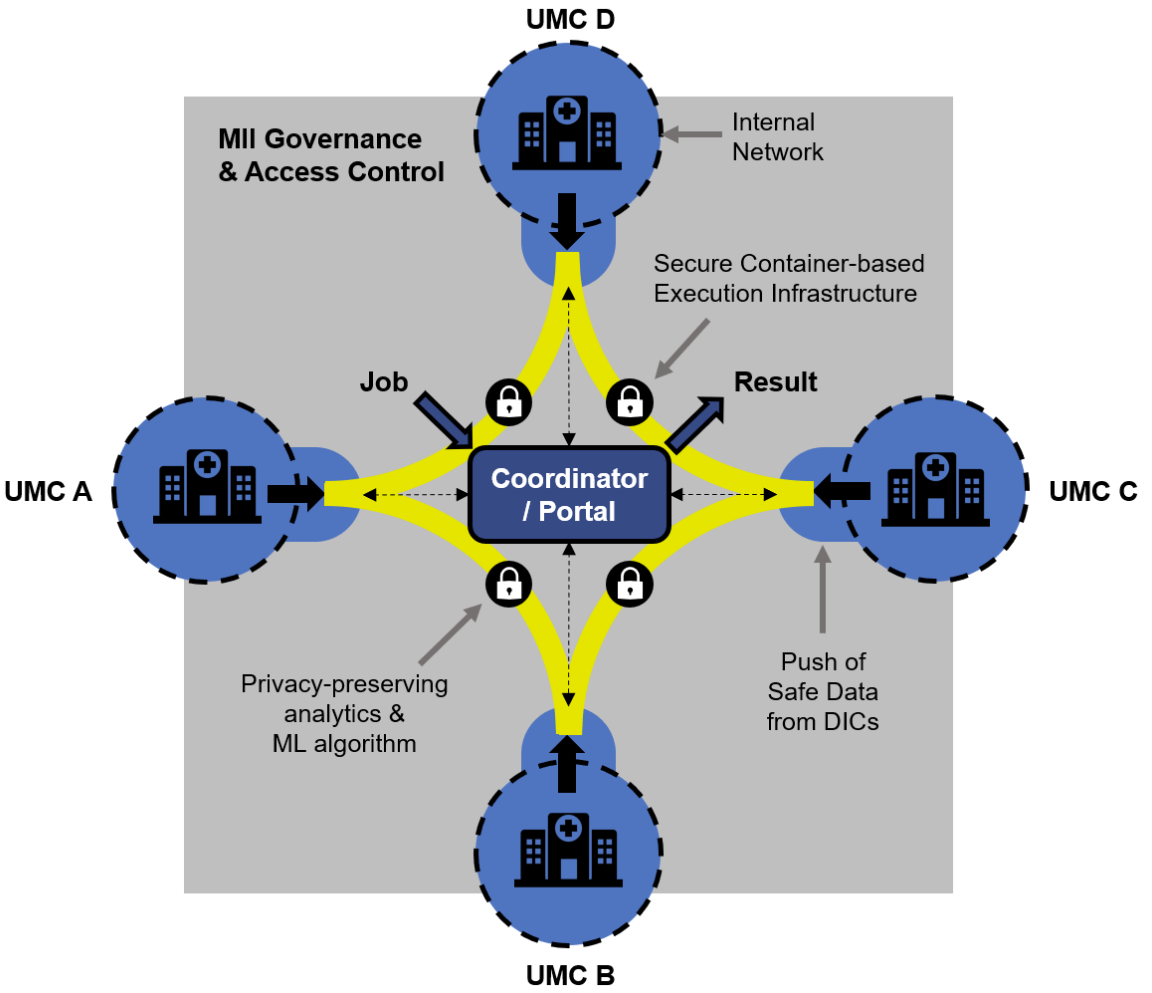
PrivateAIM - Consortium

- 15 Participants from all four MII consortia (and beyond)
- Coordinators
  - >Oliver Kohlbacher (U Tübingen)
  - >Fabian Prasser (Charité)
  - >Daniel Rückert (TU Munich)
- Three associated junior research groups
  - Mete Akgün - Medical Data Privacy and Privacy-Preserving ML on Healthcare Data (MDPPML) (Tübingen)
  - Michael Kamp – Trustworthy Machine Learning (Essen)
  - Björn Schreiweis – Medical Informatics (Kiel)

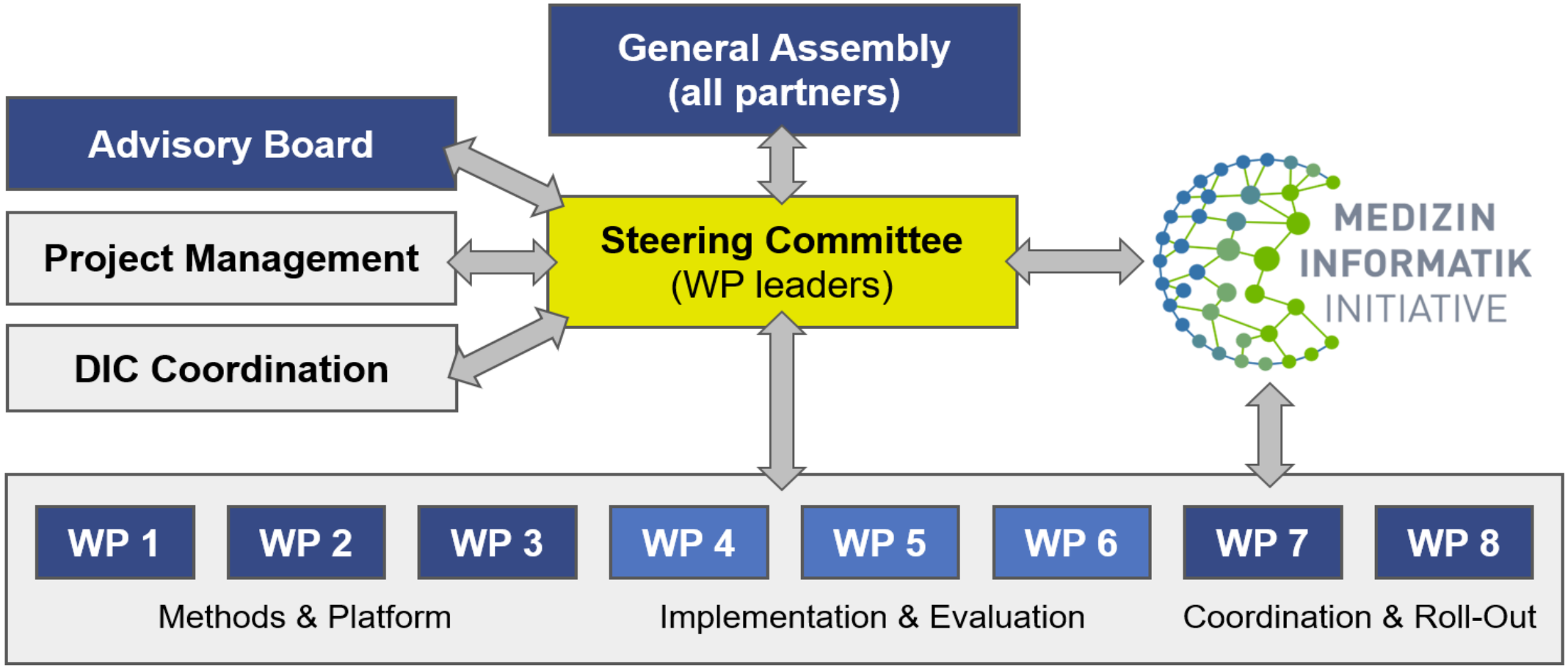
Charité - Universitätsmedizin Berlin (Charité)	Prof. Dr. Fabian Prasser
Helmholtz Center for Information Security (CISPA)	Prof. Dr. Mario Fritz
Deutsches Krebsforschungszentrum (DKFZ)	Dr. Ralf Omar Floca
University of Tübingen (EKUT)	Prof. Dr. Nico Pfeifer
Ludwig-Maximilians-Universität München (LMU)	Prof. Dr. Ulrich Mansmann
Technology, Methods, and Infrastructure for Networked Medical Research (TMF)	Dr. Sebastian C. Semler
Technische Universität München (TUM)	Prof. Dr. Daniel Rückert
Friedrich-Alexander-Universität Erlangen-Nürnberg (UKER)	Prof. Dr. Thomas Ganslandt
University of Freiburg (UKFR)	Prof. Dr. Harald Binder
University Hospital Heidelberg (UKHD)	Prof. Dr. Christoph Dieterich
University of Cologne (UKK)	Prof. Dr. Oya Beyan
Leipzig University Medical Center (UKL)	Prof. Dr. Toralf Kirsten
University Hospital Tübingen (UKT)	Prof. Dr. Oliver Kohlbacher
Ulm University (UKU)	Prof. Dr. Hans Kestler
Medical Faculty Mannheim, Heidelberg University (UMM)	Prof. Dr. Martin Lablans

PrivateAIM – Key Ideas

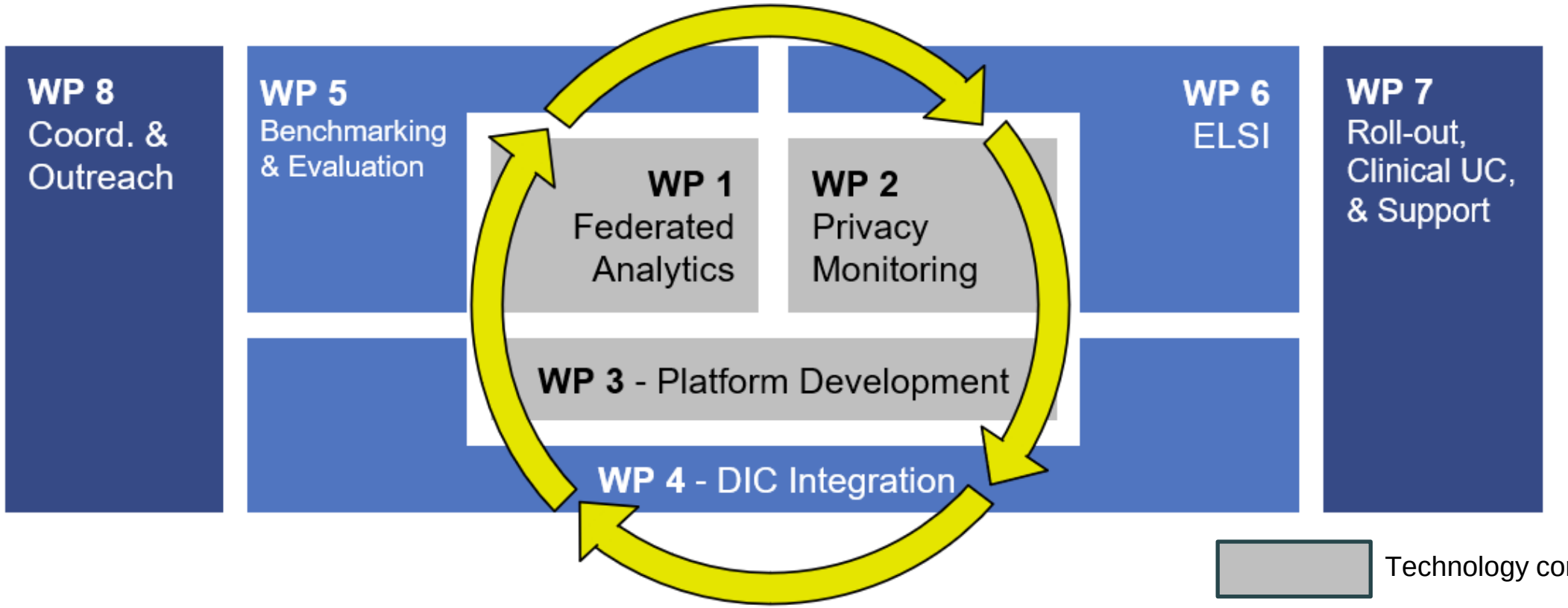
- Make major contributions in
  - Methods for federated machine learning
  - Privacy guarantees for federated analytics
  - Real-world platform for privacy-preserving analytics
- Deploy these ideas in a consistent platform across the MII sites
- Support other (clinical) use cases within the MII with the platform



PrivateAIM – Structure



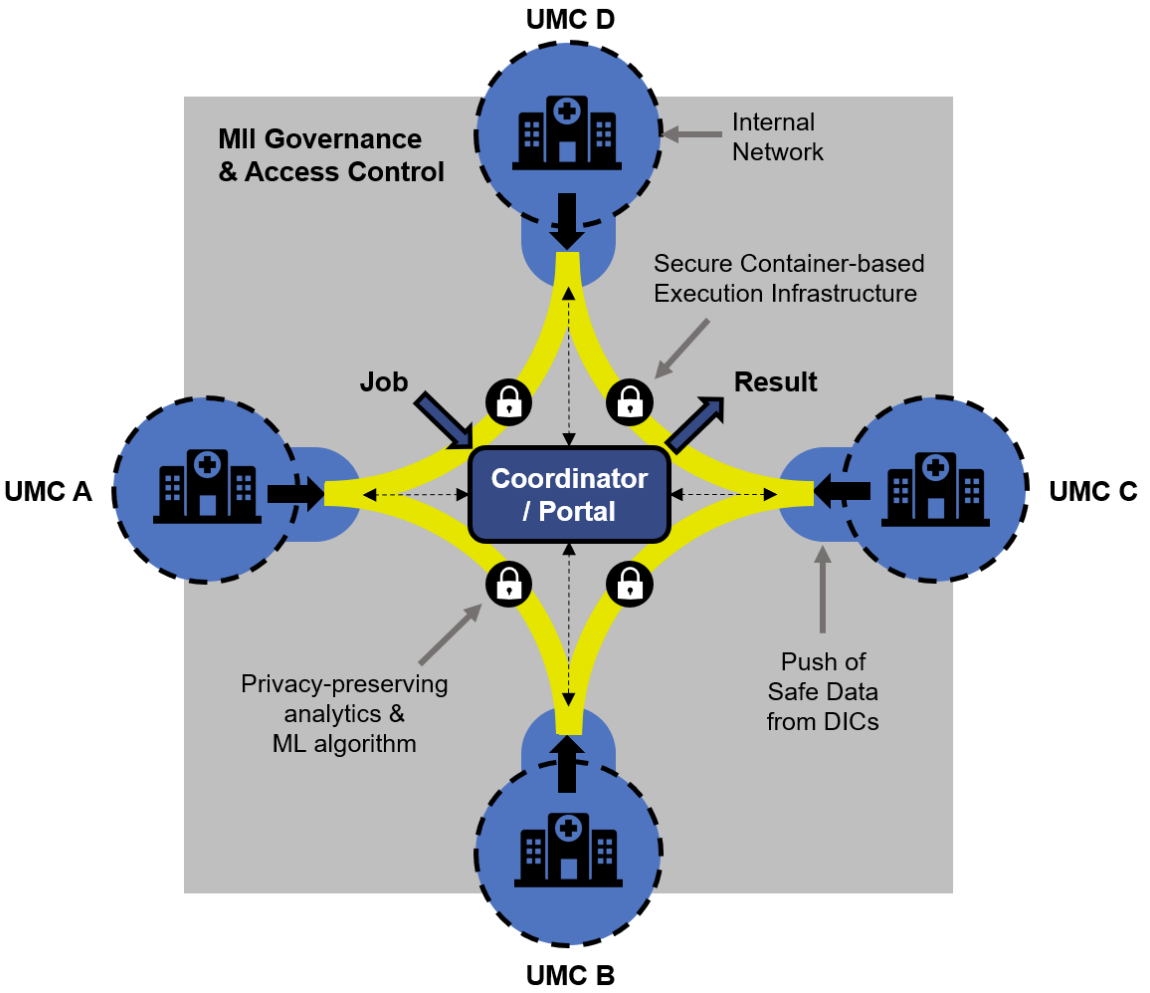
PrivateAIM – Work Packages



PrivateAIM – WP1 – Federated Analytics and ML

Federated Analytics and Machine Learning (WP1)

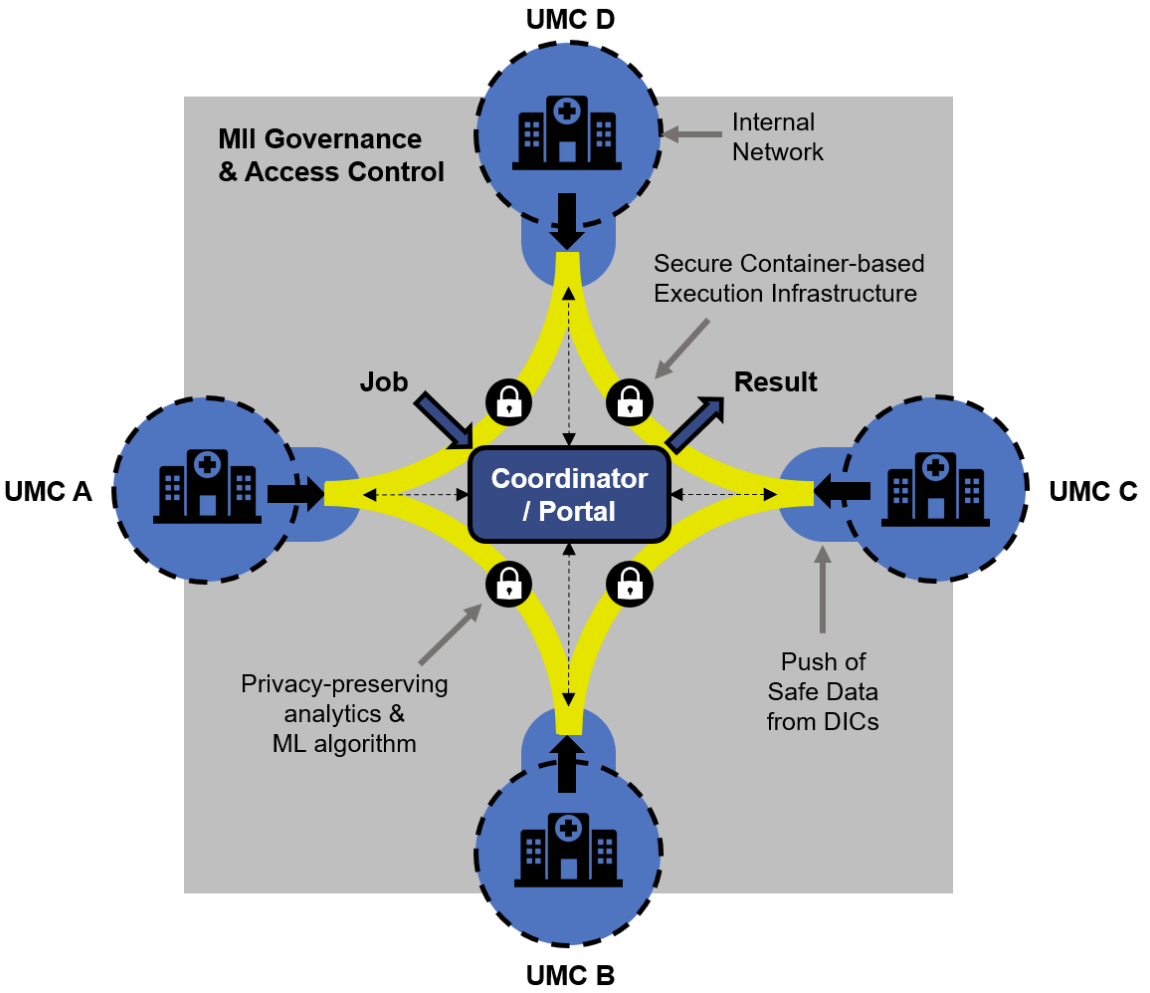
- Federated machine learning approaches that can deal with vertically and horizontally partitioned data (including non-independent and identically distributed data)
- Methods for zero/few shot learning approaches for domain adaptation/transfer learning on multi-modal medical data
- Approaches for balancing and resolving trade-offs between privacy and utility as well as privacy, fairness and robustness of ML models.



PrivateAIM – WP2 – Privacy Monitoring and Guarantees

Privacy Monitoring and Guarantees (WP2)

- Privacy frameworks combining guarantees for combinations of different types of data and multi-modal datasets.
- Translation of privacy accounting, enforcement and monitoring techniques into practical biomedical settings while maintaining utility.
- Certification of components and their privacy properties to improve trust and make transparent privacy protection under the honest-but-curious model.

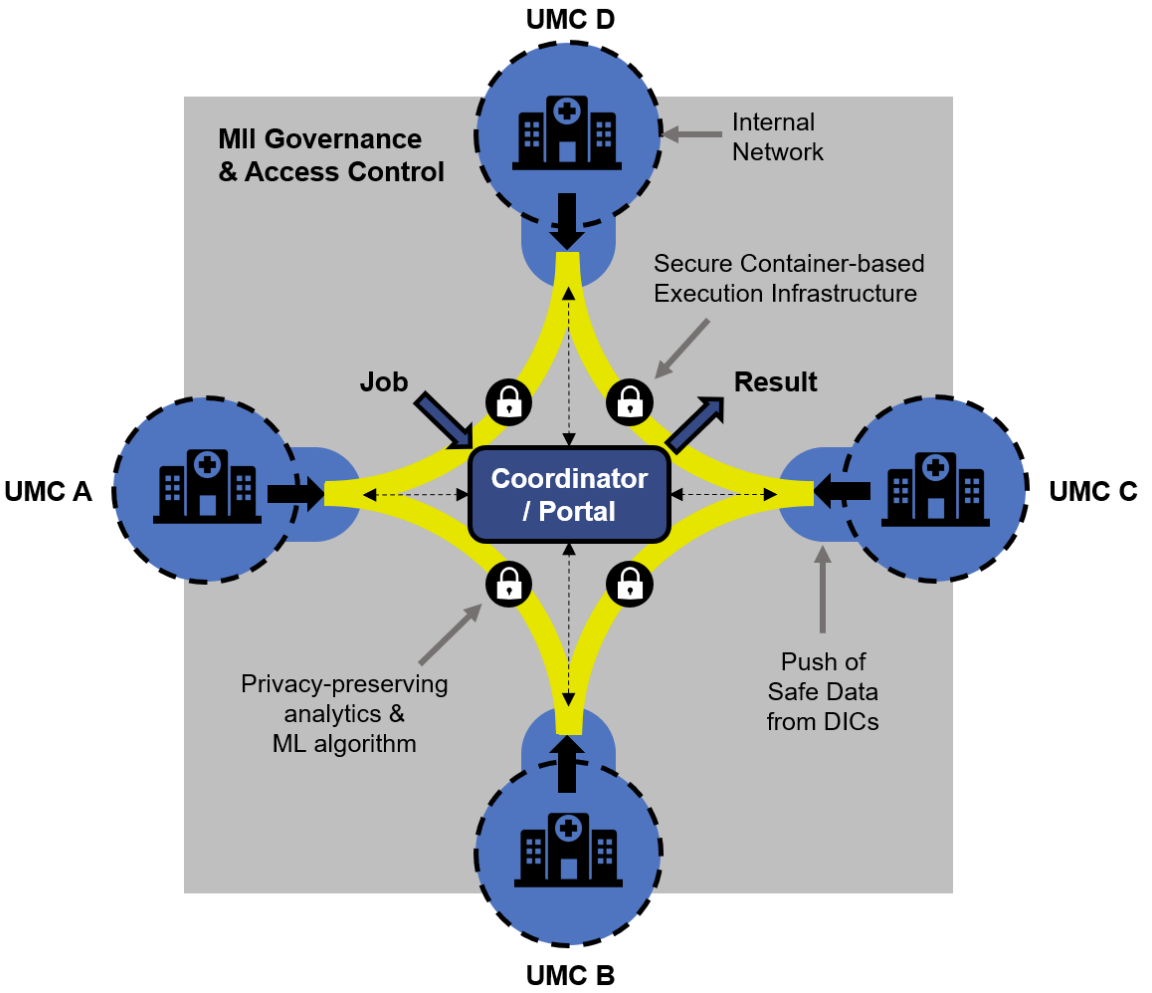


PrivateAIM – WP3 – Platform Development

Federated Analytics Platform (WP3)

- Scalable infrastructures for federated and reproducible processing of clinical, omics, and imaging data with light-weight containerized components.
- Design and integrate interoperability mechanisms with other distributed analysis platforms

The **Federated Learning and Analytics in MEdicine (FLAME) platform**, developed within this project and compiled in WP3, forms the core of the project and brings together the fundamental research questions addressed in WP1 and WP2 with the practical application within the MII in WP4.



PrivateAIM – WP4, WP5, WP6, WP7: Roll-Out and Evaluation

Deployment within the MII (WP4)

- Deployment-ready software stack and documentation for the platform to facilitate broad roll-out based on the prototypical deployment at DICs from each MII consortium
- Application of the platform to relevant data use projects within the MII.

Benchmarking and Evaluation (WP5)

- Preparation of benchmark datasets and scenariosPrivacy and utility evaluations

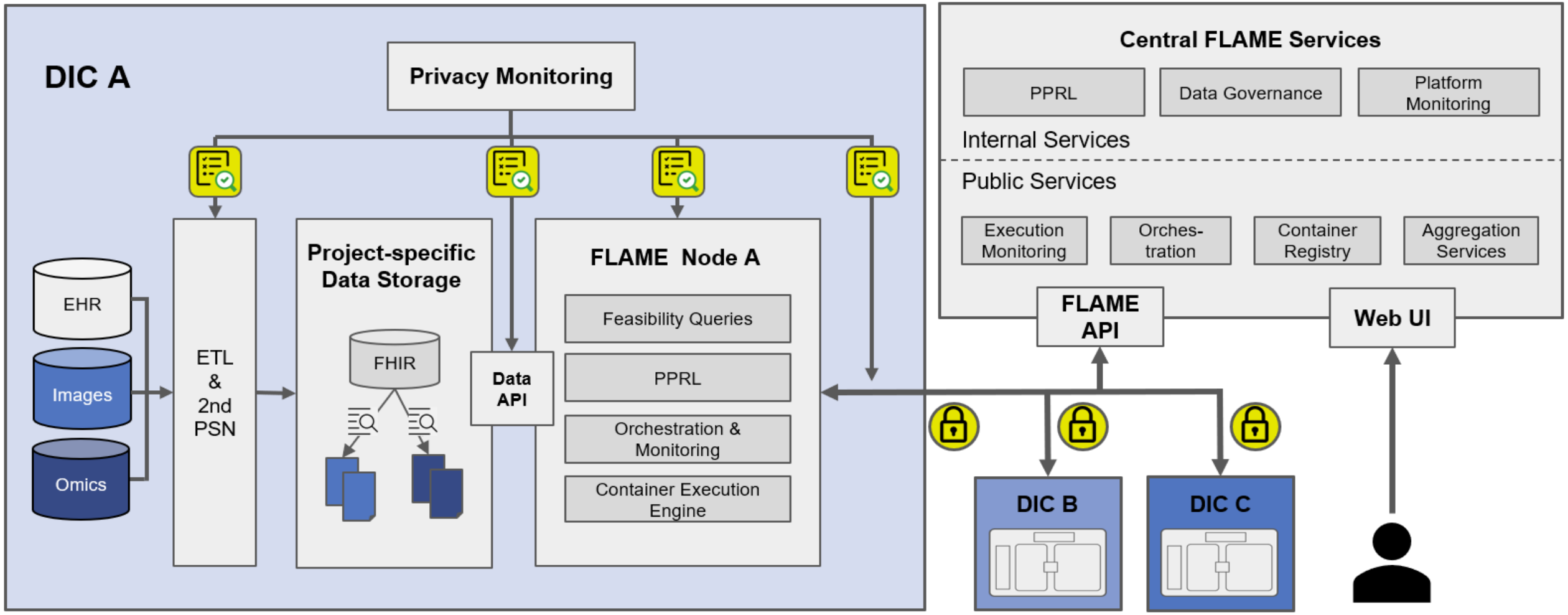
Ethical, legal and societal issues (WP6)

- Information materials for public, ethics committees, data protection officers
- Legal and regulatory framework

Roll-out (WP7)

- Other Module 3 projects and clinical use cases

The envisioned FLAME Platform



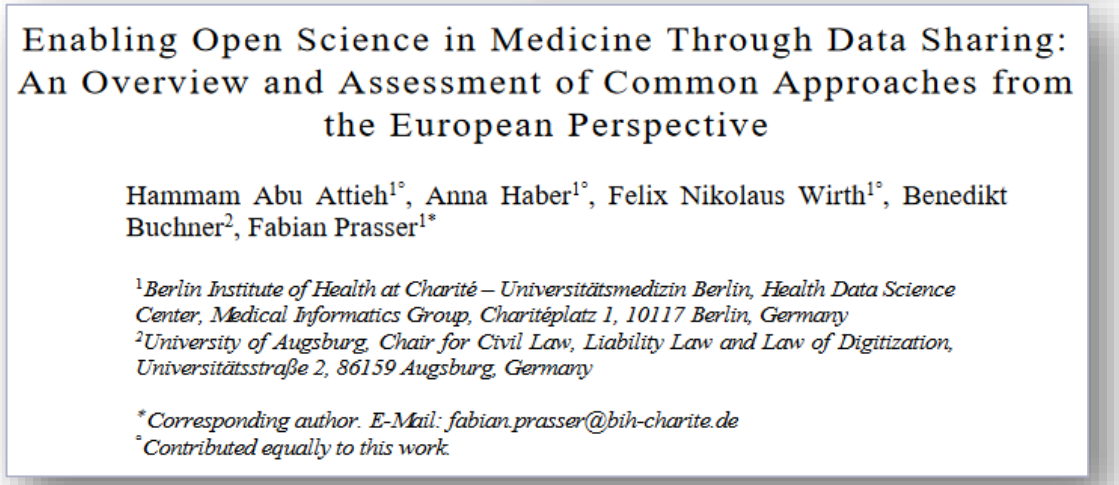
We don‘t start from scratch! Some examples (1)

RESEARCH Open Access

Privacy-preserving data sharing infrastructures for medical research: systematization and comparison

Felix Nikolaus Wirth*, Thierry Meurers, Marco Johns and Fabian Prasser

Data Sharing Architectures



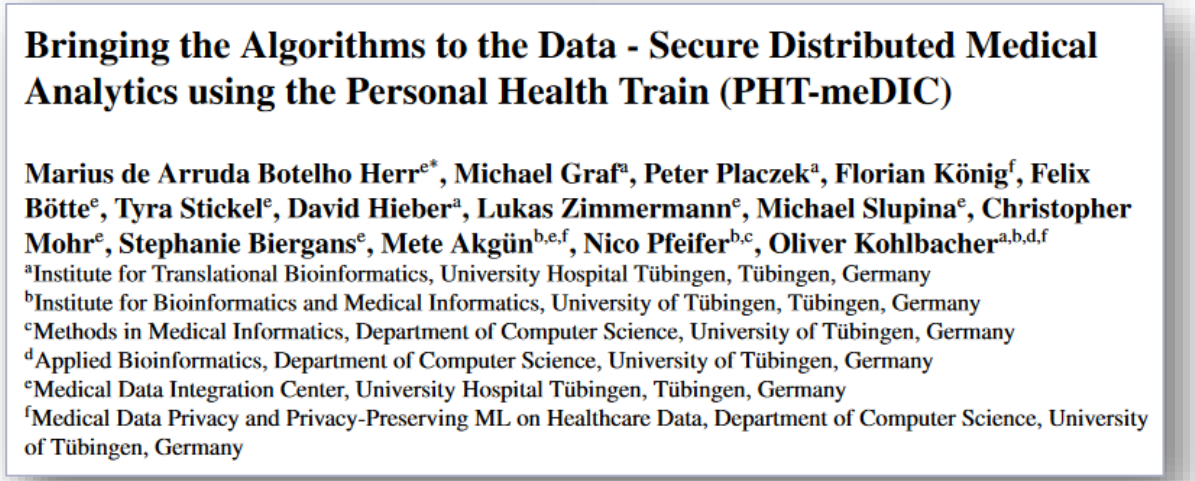
Enabling Open Science in Medicine Through Data Sharing: An Overview and Assessment of Common Approaches from the European Perspective

Hammam Abu Attieh[1°], Anna Haber[1°], Felix Nikolaus Wirth[1°], Benedikt Buchner[2], Fabian Prasser[1*]

[1]*Berlin Institute of Health at Charité – Universitätsmedizin Berlin, Health Data Science Center, Medical Informatics Group, Charitéplatz 1, 10117 Berlin, Germany*
[2]*University of Augsburg, Chair for Civil Law, Liability Law and Law of Digitization, Universitätsstraße 2, 86159 Augsburg, Germany*

*Corresponding author. E-Mail: fabian.prasser@bih-charite.de
°Contributed equally to this work.

Technico-Legal Analyses

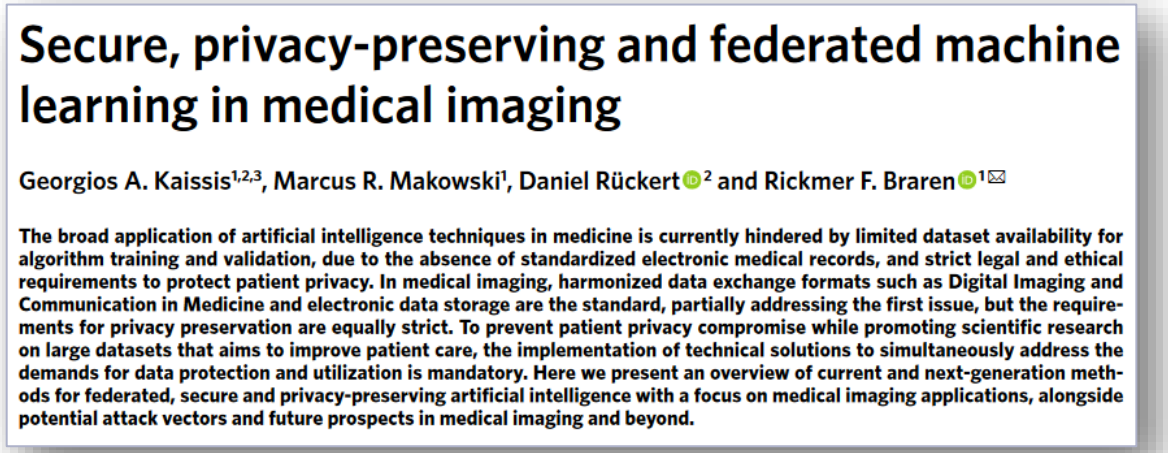


Bringing the Algorithms to the Data - Secure Distributed Medical Analytics using the Personal Health Train (PHT-meDIC)

Marius de Arruda Botelho Herr[e*], Michael Graf[a], Peter Placzek[a], Florian König[f], Felix Bötte[e], Tyra Stickel[e], David Hieber[a], Lukas Zimmermann[e], Michael Slupina[e], Christopher Mohr[e], Stephanie Biergans[e], Mete Akgün[b,e,f], Nico Pfeifer[b,c], Oliver Kohlbacher[a,b,d,f]
[a]Institute for Translational Bioinformatics, University Hospital Tübingen, Tübingen, Germany
[b]Institute for Bioinformatics and Medical Informatics, University of Tübingen, Tübingen, Germany
[c]Methods in Medical Informatics, Department of Computer Science, University of Tübingen, Germany
[d]Applied Bioinformatics, Department of Computer Science, University of Tübingen, Germany
[e]Medical Data Integration Center, University Hospital Tübingen, Tübingen, Germany
[f]Medical Data Privacy and Privacy-Preserving ML on Healthcare Data, Department of Computer Science, University of Tübingen, Germany

PHT Implementation

We don‘t start from scratch! Some examples (2)



Secure, privacy-preserving and federated machine learning in medical imaging

Georgios A. Kaissis[1,2,3], Marcus R. Makowski[1], Daniel Rückert[2] and Rickmer F. Braren[1]✉

The broad application of artificial intelligence techniques in medicine is currently hindered by limited dataset availability for algorithm training and validation, due to the absence of standardized electronic medical records, and strict legal and ethical requirements to protect patient privacy. In medical imaging, harmonized data exchange formats such as Digital Imaging and Communication in Medicine and electronic data storage are the standard, partially addressing the first issue, but the requirements for privacy preservation are equally strict. To prevent patient privacy compromise while promoting scientific research on large datasets that aims to improve patient care, the implementation of technical solutions to simultaneously address the demands for data protection and utilization is mandatory. Here we present an overview of current and next-generation methods for federated, secure and privacy-preserving artificial intelligence with a focus on medical imaging applications, alongside potential attack vectors and future prospects in medical imaging and beyond.

Privacy-Preserving Machine Learning

TECHNICAL NOTE

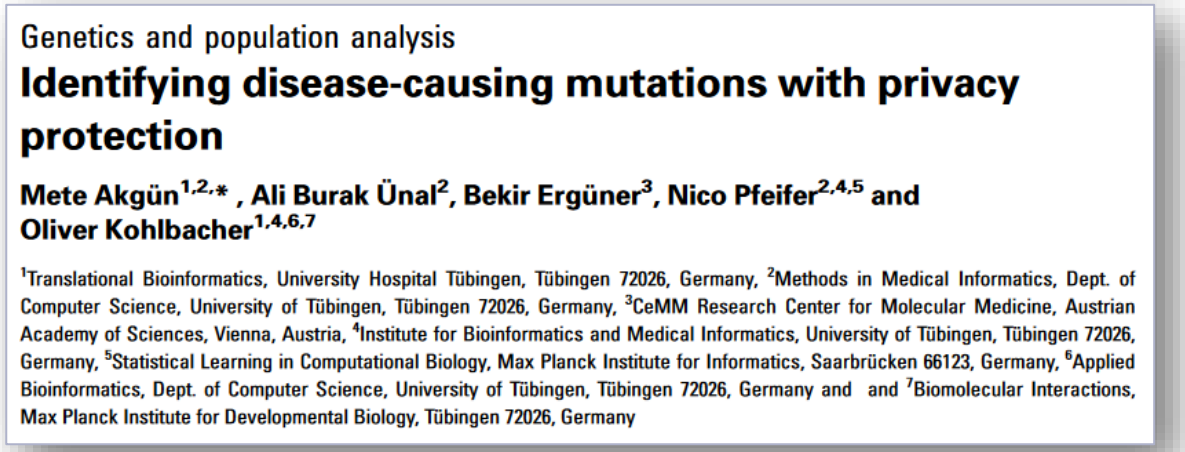
A scalable software solution for anonymizing high-dimensional biomedical data

Thierry Meurers[1,*], Raffael Bild[2], Kieu-Mi Do[3] and Fabian Prasser[1]

[1]Berlin Institute of Health at Charité–Universitätsmedizin Berlin, Medical Informatics, Charitéplatz 1, 10117 Berlin, Germany; [2]School of Medicine, Technical University of Munich, Ismaninger Str. 22, 81675 Munich, Germany and [3]Faculty of Informatics, Technical University of Munich, Boltzmannstr. 3, 85748 Garching, Germany

*Correspondence address. Thierry Meurers, Berlin Institute of Health at Charité–Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany. E-mail: thierry.meurers@charite.de https://orcid.org/0000-0001-8168-7067

Data Anonymization



Genetics and population analysis

Identifying disease-causing mutations with privacy protection

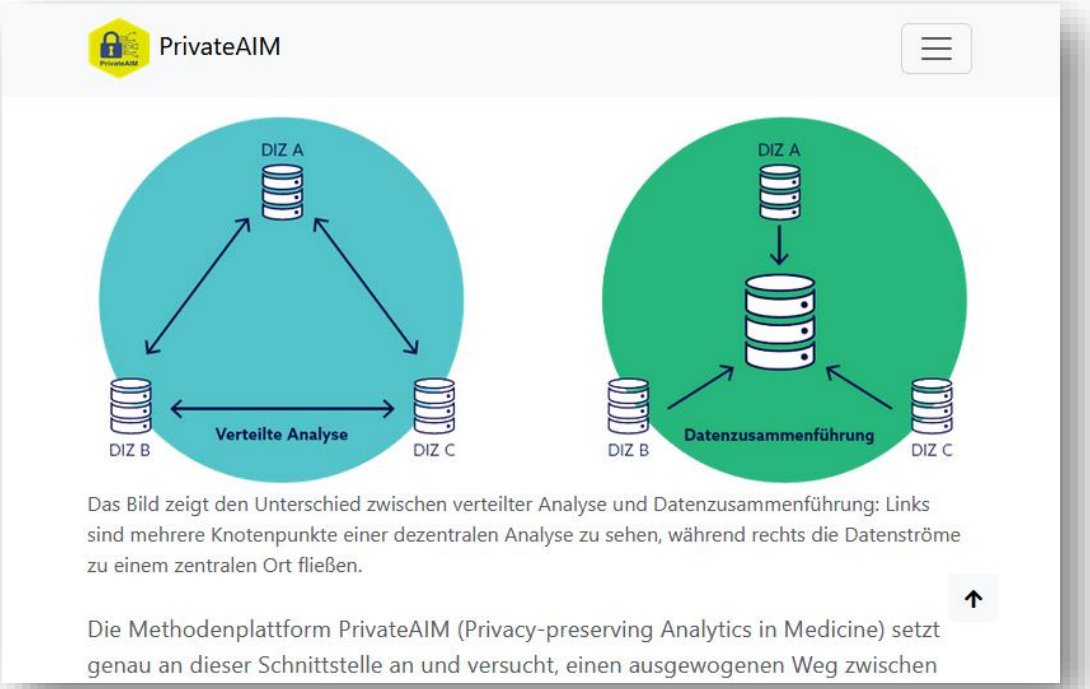
Mete Akgün[1,2,*], Ali Burak Ünal[2], Bekir Ergüner[3], Nico Pfeifer[2,4,5] and Oliver Kohlbacher[1,4,6,7]

[1]Translational Bioinformatics, University Hospital Tübingen, Tübingen 72026, Germany, [2]Methods in Medical Informatics, Dept. of Computer Science, University of Tübingen, Tübingen 72026, Germany, [3]CeMM Research Center for Molecular Medicine, Austrian Academy of Sciences, Vienna, Austria, [4]Institute for Bioinformatics and Medical Informatics, University of Tübingen, Tübingen 72026, Germany, [5]Statistical Learning in Computational Biology, Max Planck Institute for Informatics, Saarbrücken 66123, Germany, [6]Applied Bioinformatics, Dept. of Computer Science, University of Tübingen, Tübingen 72026, Germany and and [7]Biomolecular Interactions, Max Planck Institute for Developmental Biology, Tübingen 72026, Germany

Secure Multi-Party Computation

Outlook

- PrivateAIM project started in April 2023 as one of the first method platforms
- We have established a strong interdisciplinary team working on all aspects of federated privacy-preserving analytics
- The FLAME platform developed within the project will be the successor of the PHT infrastructure currently deployed within some MII consortia
- We expect PrivateAIM to build the foundation for federated AI within the MII and support other clinical use cases



Visit us at: https://privateaim.de

Thank you!

GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung