

# IT Infrastructure Overview of the RWTH cBMB

Philipp Leusmann<sup>1</sup>, Jürgen Veeck<sup>1</sup>, Jörg Jäkel<sup>1</sup>, Cord Spreckelsen<sup>2</sup>,  
 Ruth Knüchel-Clarke<sup>1</sup> and Edgar Dahl<sup>1</sup>

<sup>1</sup>Institute of Pathology; <sup>2</sup>Institute of Medical Informatics,  
 University Hospital of RWTH Aachen

## Environment

Beside the handling and physical storage of biomaterial samples, a properly run Biomaterial Bank requires a sound IT infrastructure.

The RWTH cBMB has formed a dedicated IT team to accomplish all data related challenges. To provide the technical preconditions for this, the RWTH cBMB is closely cooperating also with the IT team of the University Hospital. All provided services are run on machines administrated by them. Additionally, a firewall and backup facilities are provided.

## Data Distribution

Data maintained in the RWTH cBMB is divided into four context dependent categories:

- Organizational data (OrgDAT)
- Personally Identifying data (IDAT)
- Medical data (MDAT)
- Sample Analytical data (ProbDAT)

These data types have been determined by the TMF in collaboration with the data privacy officers of the German federal states. For meeting the legislative requirements, three of them are stored distributed over several systems. Storing ProbDAT is currently not implemented.

## Organizational Data

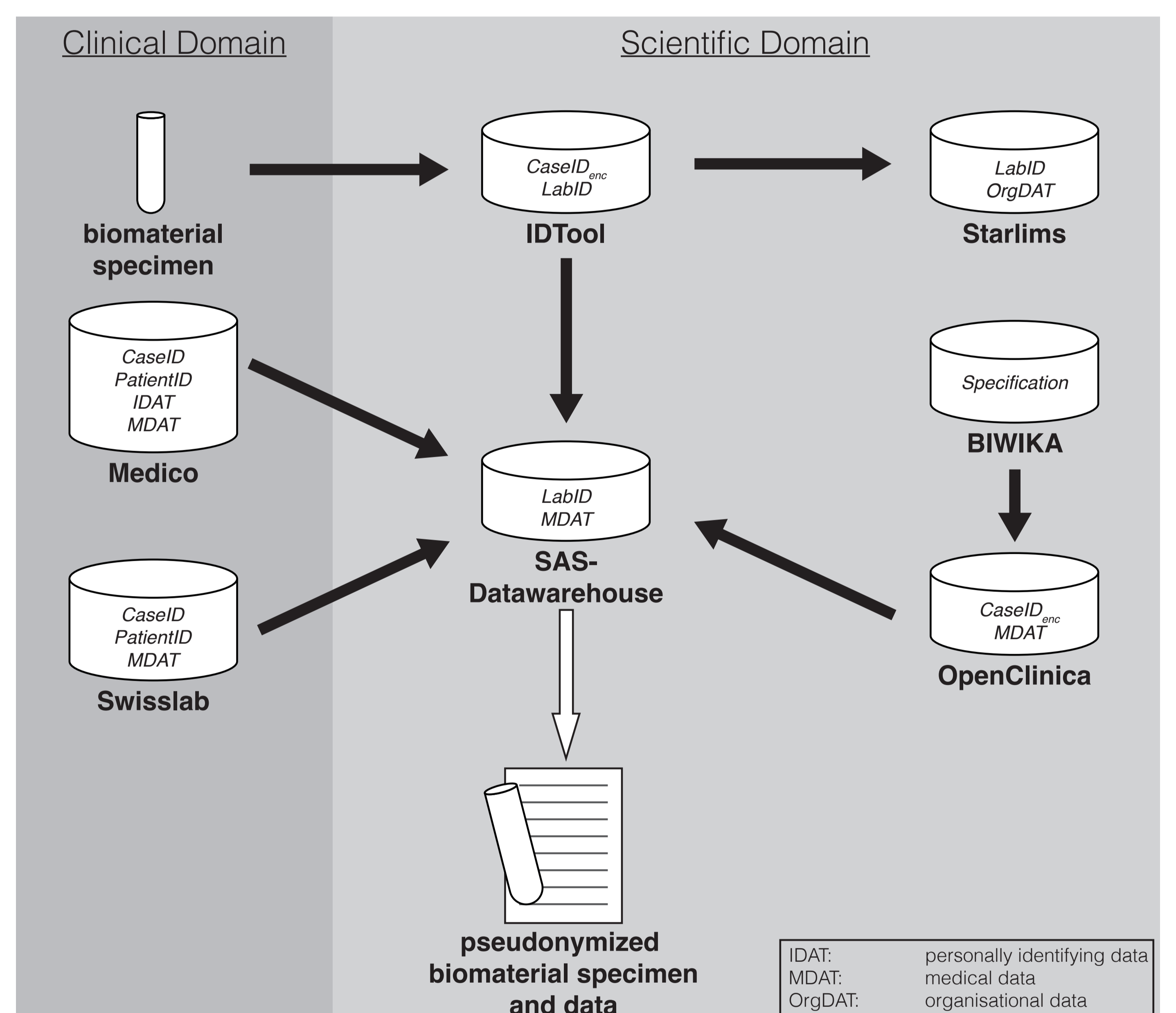
To maintain the organizational data of a stored sample, it is registered in the Starlims Laboratory Information Management System. For each sample properties such as location, sending institute, receipt and extraction time, processing characteristics and further sample or project dependent properties are collected.

## Personally Identifying Data

The main repository for patient related personal data in the University Hospital Aachen is the clinical information system MEDICO. As access to this system is only provided to physicians in the context of treatment, the stored data is not available directly for use in the RWTH cBMB. Since most parts of the personal data are irrelevant for research purposes, the existing data solely remains in this system.

## Medical Data

Medical data related to a treatment and thus to a sample is in most cases collected during clinical routine operation. Consequently, it is stored in the corresponding clinical systems which are MEDICO for the general case or the Swisslab system in the case of the Institute of Pathology. Alternatively, samples may be collected specifically for a certain research project. In this case, sophisticated project related medical parameters need to be acquired and stored in a dedicated system using a pseudonym. The RWTH cBMB plans to provide an OpenClinica server for this purpose. The design of the necessary project specification is supported by the collaborative approach of the BIWIKA semantic wiki system.



The IT systems at the RWTH cBMB with indication of data flows.

## Linking Distributed Data

The self-developed IDTool provides a link between the pseudonymized research data, the samples and the personally identifying data. The use of asymmetric encryption ensures the protection against unauthorized disclosure of personally identifying data.

Using the publicly available encrypted case or sample identifier, a pseudonymized combined view of the distributed data linked to a sample can be made available by the SAS-Datawarehouse.