centralized Biomaterial Bank
RWTH AACHEN UNIVERSITY

IDTool
Pseudonymization service development

Alexander Devaykin[1], Philipp Leusmann[1], Jürgen Veeck[1], Jörg Jäkel[1], Cord Spreckelsen[2], Ruth Knüchel-Clarke[1] and Edgar Dahl[1]

[1]Institute of Pathology; [2]Institute of Medical Informatics, University Hospital of the RWTH Aachen

GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

UNIVERSITÄTSKLINIKUM AACHEN

IDTool

IDTool is a web application developed by RWTH cBMB for storing biomaterial identification data according to data privacy law. This tool provides the foundation for linking pseudonymized samples to the personally identifying data of the donor.

The IDTool generates random unique identifiers for submitted samples and further identification properties, such as patient or case identifiers used in the clinical management systems. These entities are linked in a way that any information able to disclose personally identifying data is protected.

IDTool key features are:

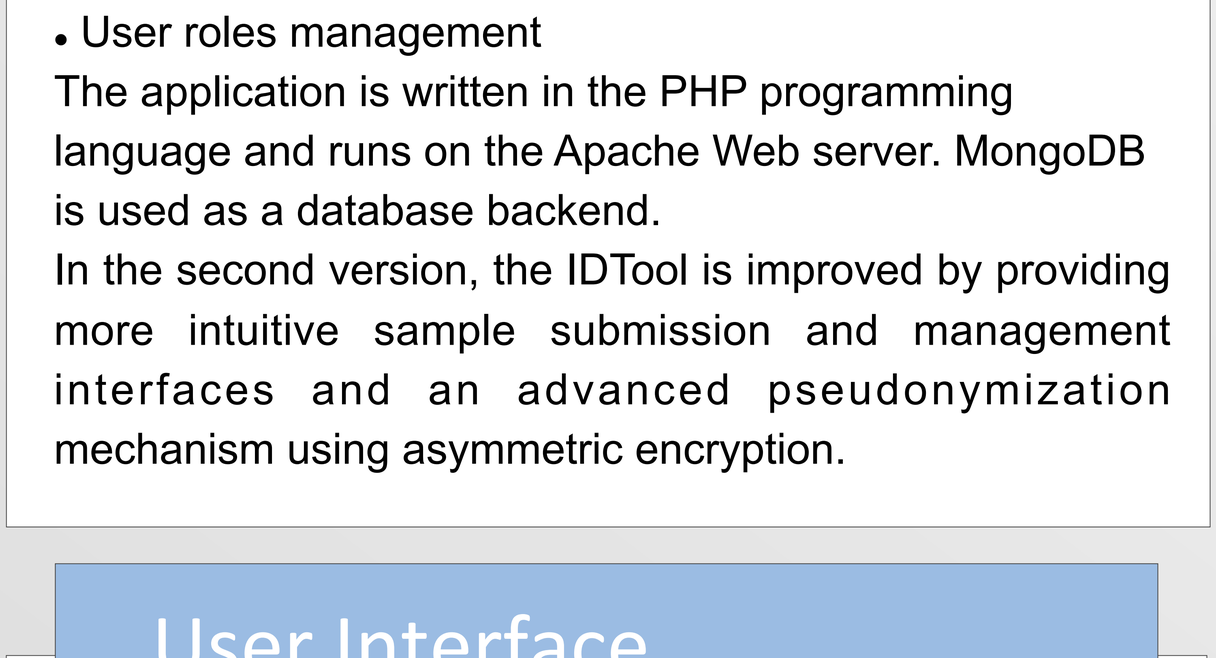
- Samples submission, search and management
- Printing labels for sample containers
- Tracking of change history
- User roles management

The application is written in the PHP programming language and runs on the Apache Web server. MongoDB is used as a database backend.

In the second version, the IDTool is improved by providing more intuitive sample submission and management interfaces and an advanced pseudonymization mechanism using asymmetric encryption.

User Interface

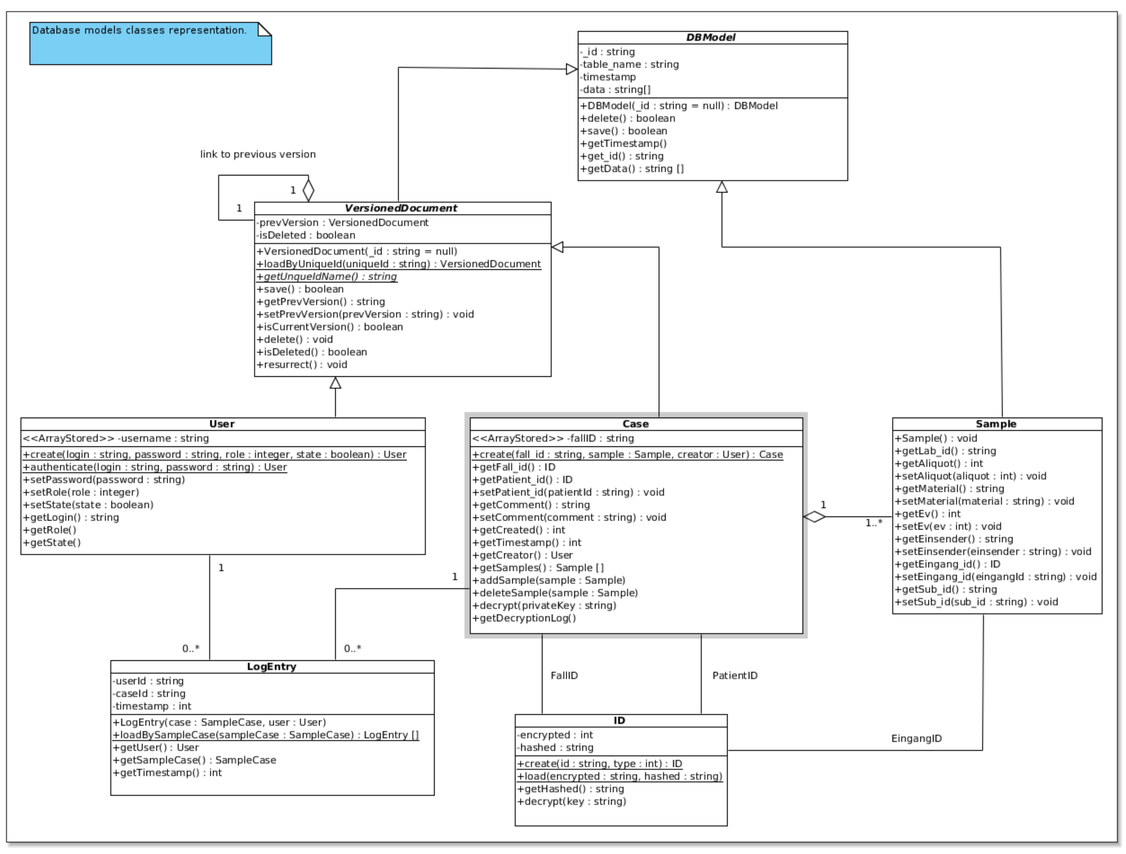
The user interface has been designed according to laboratory personnel workflow. It facilitates easy access to the most frequently used functions. The sample submission interface has been optimized in regard to avoidance of redundant data input.



Example of the user interface (development mockup): user does not need to input case data for a sample, if the case already exists in the system.

Modelling the System

Modern software engineering approaches have been used for the IDTool modeling on all stages of the development process. UML diagrams have been used for modeling user roles management, sample submission and editing, database structure and relationships. Several diagram types, including class, sequence, activity and ER diagrams have been created.



The class diagram representing database documents. *DBModell* class is inherited by all database documents, *VersionedDocument* class is inherited by the documents with the need of history tracking.

Security

The encryption model has been significantly improved in the second major IDTool version. Important sample data is encrypted using the RSA asymmetric encryption algorithm with 2048-bit key length. This way only the private key owner (data trustee) can decrypt data, while the public key is used for encryption. This approach avoids the necessity to store private keys on the server and allows to effectively log all decryption attempts. As the resulting encrypted value is too large to be used as a pseudonym, hashes are generated additionally to be used as identifiers in place of the encrypted value. These identifiers also contain a checksum to allow validation during data entry from the cooperating institutions. The resulting pseudonyms are 12 character ASCII strings, so they stay readable and easy to input.