Bringing heterogenous Research Data together: Data Protection for the German Centre for Lung Research (DZL)

Raphael W. MAJEEDa,1, Stefan KUHNa, Clemens RUPPERTa, Andreas GÜNThERA, and Rainer RÖHRIG a

Justus-Liebig-University Giessen, Germany

Keywords. Research Network, Lung, Database, Data Protection, Data Warehouse

Introduction

The German Centre for Lung Research (DZL) is an association of Germany’s leading university and non-university institutions and aims to foster translational lung research. The objective of our work was to develop a software architecture providing a central data repository to decentralised research network members, while adhering to the complex German data protection directives. The resulting software will be used to link content from different existing research databases, thus producing a comprehensive DZL-wide database for lung related research.

1. Methods

Anonymisation of patient-identifying information is achieved using deterministic algorithms from asymmetric public-key-cryptography. Re-identification risk is minimized by employing data transformation to comply with k-anonymity, l-diversity and t-closeness.

2. Results

Local sites submit pseudonyms and encrypted medical data to the identity management service. Pseudonyms are translated and mapped. Medical data is forwarded to the data warehouse. Researchers can perform queries and order data and specimen.

3. Discussion

Data protection regulations in Germany are among the strictest in the world. Thus, the presented architecture should be applicable to other countries as well.

1 Corresponding author: raphael.majeed@chiru.med.uni-giessen.de