The Leipzig Health Atlas

A Repository for Research Data and Models

C. Beger1, R. Hänsel1, F. Meineke1, S. Stäubert1, M. Löbe1, A. Uciteli1, H. Binder2, T. Kirsten3,4, M. Scholz1, H. Herre1, C. Engel1, M. Löfler1,2,3,5

1 University of Leipzig, Institute for Medical Informatics, Statistics and Epidemiology (IMISE), Leipzig, Germany
2 University of Leipzig, Interdisciplinary Centre for Bioinformatics (IZBI), Leipzig, Germany
3 University of Leipzig, Research Centre for Civilization Diseases (LIFE), Leipzig, Germany
4 University of Leipzig, Faculty of Applied Computer and Biological Sciences, Mittweida, Germany
5 University of Leipzig, Clinical Trial Centre Leipzig (ZKS Leipzig), Leipzig, Germany

The Leipzig Health Atlas (LHA) is a platform for presentation and exchange of publications, (bio-) medical data, models and software tools from the field of health research. The LHA is based on the FAIR Data Principles (Findable, Accessible, Interoperable, Reusable). Currently, more than 18 research consortia contribute data from the domains lymphoma, glioma, sepsis, hereditary colorectal and breast cancer, and others. Target groups are clinicians, epidemiologists, molecular geneticists, human geneticists, pathologists, biostatisticians and modellers. The LHA is - under consideration of legal access restrictions - open to the public at www.health-atlas.de.

The components of the Leipzig Health Atlas (LHA) are managed as Docker containers via a Kubernetes infrastructure. For all data sets, the underlying diseases are annotated using Human Disease Ontology. This facilitates navigation and allows statements to be made about the scope and coverage of the repository.

The Leipzig Health Atlas manages information on projects, people involved, data sets, models and publications in a permanent archive. The assigned identifiers can be referenced and resolved externally.

Projects in the fields of bioinformatics, biomath and computational biology process large, heterogeneous data in complex pipelines that have to be published for the reproducibility and reusability of results.

Parametrizable models are made available in the R programming language as interactive Shiny Apps.

Tools for browsing through data sets allow the estimation of available data for new research hypotheses (Figure: JKB).

With graphical analysis tools (e.g., tranSMART), statistical evaluations such as survival curves, correlations, p-values or regression lines can be generated with a mouse click.

As an example, an evaluation of the strength of the subjects’ handgrip on a point scale and the geographical distribution over the districts of the city of Leipzig (after clicking on “Try App”).

The Leipzig Health Atlas offers a platform for presentation and exchange of publications, (bio-) medical data, models and software tools from the field of health research. The LHA is based on the FAIR Data Principles (Findable, Accessible, Interoperable, Reusable). Currently, more than 18 research consortia contribute data from the domains lymphoma, glioma, sepsis, hereditary colorectal and breast cancer, and others. Target groups are clinicians, epidemiologists, molecular geneticists, human geneticists, pathologists, biostatisticians and modellers. The LHA is - under consideration of legal access restrictions - open to the public at www.health-atlas.de.

For more information, please visit www.health-atlas.de.