GERMAN NETWORK FOR BIOINFORMATICS INFRASTRUCTURE

Virtual 3rd Annual Meeting of the de.NBI Industrial Forum

Book of Abstracts – Scientific Presentations

November 24th 2022, 01:00-4:30 pm (CET)

Virtual Meeting using Zoom

Log-In Data will be send to you after the registration

Section B - Digital Health activities of different organizations

How do ELIXIR's bioinformatic infrastructures contribute to national and European digital health innovative actions?

Despoina Sousoni, Industry Officer, ELIXIR Hub, Hinxton, England

Abstract:

ELIXIR is an intergovernmental organisation that brings together life science resources across Europe, enabling users in academia and industry to access services vital for their research. This access to resources supports the creation of an open innovation ecosystem in the life sciences in Europe. ELIXIR Nodes have always played a key role in national and international data-driven health initiatives by providing unique bioinformatic resources and engaging with industry. Examples of ELIXIR academia-industry engagement in the digital health sector include a bottom-up effort to organise national bioinformatics activities and align them with the needs of the private sector coordinated by ELIXIR Spain, a personalised health research infrastructure with ELIXIR Netherlands which includes industry as a partner, and involvement of ELIXIR Finland and ELIXIR Sweden in a European Innovative Health Initiative project to build a central repository for digital pathology. These are some best practices of how academia and industry can collaborate to establish robust infrastructures at the national and European levels.

NFDI4Health – The National Research Data Infrastructure for Personal Health Data in the context of the European Health Data Space

Juliane Fluck, NFDI4Health, ZB MED Cologne

Abstract:

Within the framework of the NFDI, valuable data resources from science and research are to be systematically accessed, networked and made usable in a sustainable and qualitative manner for the entire German science system. The goals of NFDI4Health include improving the discoverability of and access to clinical, epidemiological and public health studies conducted in Germany and enabling the exchange and linking of these data to other health data while ensuring data protection.

To enable discoverability, we have developed a general metadata schema for study data. This scheme is currently being tested and validated with other health infrastructure projects on a national and international level. The goal is to create an internationally accepted metadata standard that will enable easy exchange of published metadata. In addition, NFDI4Health aims to increase findability through an easy-to-use, harmonized search and exploration platform. The German Central Health Study Hub COVID-19 (https://covid19.studyhub.nfdi4health.de), developed by the Task Force COVID-19, is such a service and currently bundles over 1,250 (Sept. 2022) epidemiological, public health, and clinical studies into a single data repository. In addition to study characteristics, the platform also provides easy access to survey instruments. In the next version, the platform will cover further data collections in addition to COVID-19 data. Furthermore, NFDI4Health is currently setting up local data hubs that will enable distributed data analysis for selected data sets.

NFDI4Health aims to become part of the EHDS by making information on German study data discoverable, by enabling centralized and transparent data access requests in the near future, by providing platforms for distributed data analysis, and by working on better ways to link data sets in a broad consensus with German stakeholders.

From data to knowledge: Leveraging health record and omics data for prevention and clinical care

Roland Eils, Center for Digital Health, Berlin Institute of Health (BIH) at Charité - Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, Berlin, Germany

and

Health Data Science Unit, BioQuant, Medical Faculty of University of Heidelberg, Germany

Abstract:

In primary prevention, risk stratification is critical for the early identification of high-risk individuals. Beyond established clinical predictors, new biomarkers are regularly considered for clinical adoption. However, adoption into clinical practice not only requires a high additional prognostic value, but also practicability and affordability of predictors. Here, I will report on our ongoing efforts to integrate data from clinical care, health record and research data including omics profiling at scale for disease risk stratification. Along two examples I will demonstrate the potential of such data integration efforts for almost all common conditions including metabolic, vascular, respiratory, musculoskeletal and neurological diseases and cancers.

Knowledge-based machine learning from multi-omics data for personalized medicine

Julio Saez-Rodriguez, University of Heidelberg

Abstract:

Modern technologies allow us to profile in high detail biological and medical samples at fast decreasing costs. New technologies are opening new data modalities, including to measure at the single-cell level and with spatial resolution. Computational models, in particular those built with statistics and machine learning, are expected to help us to extract insight form these data. Using biological knowledge to aid machine learning can significantly improve the results. Towards this end, we have developed a number of tools that range from a meta-resource of biological knowledge to methods to infer pathway and transcription factor activities from gene expression and subsequently infer causal paths among them. Furthermore, and to complement a large-scale basal profiling of samples, we have developed approaches to build dynamic logic models of molecular networks and how they response to perturbations such as drug treatment. I will illustrate their utility in cases of biomedical relevance and show how they improve our understanding of molecular processes, identify biomarkers, and point at novel therapeutic opportunities.

Multi-Omics and Spatial Single Cell Methods in Lymphoid Malignancies

Wolfgang Huber, EMBL Heidelberg

Abstract:

While the arsenal of anti-tumor therapies is expanding (targeted small compounds, bispecific antibodies, CAR T cells), responses to therapy can vary highly from patient to patient. Our research aims at contributing to prediction of response for optimal therapy selection, as well as at molecular understanding of resistances, for possible exploitation in improving treatments e.g. by combination.

To this end, in collaboration with clinical partners we employ systematic multi-omics characterisation of large patient-derived lymphoid tumor cohorts, ex-vivo drug response assays, laboratory models of microenvironmental interactions at different levels of complexity, single-cell sequencing of transcriptomes, and spatial single cell omics of tumors in their native lymph node microenvironment. We use various statistical approaches of multi-omics data integration, including multi-omics factor analysis (MOFA), to discover prognostic markers suitable for follow-up clinical study, and mechanistic hypotheses suitable for molecular biological follow-up. I will report on recent examples in CLL and B-NHL, which exemplify feasibility and success of the approach.

Section C - **Digital Health activities of selected companies of the de.NBI Industrial Forum**

Literature based reasoning for precision medicine in hematology

Thomas Wollmann, Alexander Fecke Merantix Momentum (Merantix Labs GmbH), Berlin

Abstract:

An extensively investigated disease in the field of hematology is multiple myeloma. Effective treatment of multiple myeloma requires individualized therapy. Currently, an increasing amount of data is being produced by an increasing number of clinical studies and case reports. However, practitioners at all experience levels face information overload for individualized therapy selection based on available clinical evidence. Leipzig University, Janssen, and Merantix Momentum are developing a model-based clinical decision support (CDS) system which integrates patient data from clinical information systems with the available knowledge base. Our subsystem augments the CDS with latest patient relevant studies from PubMed by querying the literature based on a digital patient model. The approach has the potential to spark novel and personalized multiple myeloma treatment strategies, making cancer treatment more patient-specific, targeted, and effective.

Curated single-cell RNA-Seq for ML and AI applications in pre-clinical biomarker discovery and validation in the immune system

Jana Sponarova, Pavel Honsa, Klara Ruppova, and Philip Zimmermann. Nebion AG – an Immunai Company, Zurich

Abstract:

The immune system is a complex system of various cell populations and soluble mediators, and its proper understanding requires detailed analysis. These efforts have been boosted by the development of single-cell RNA sequencing (scRNAseq). ScRNAseq is a powerful approach to understanding molecular mechanisms of development and disease and uncovering cellular heterogeneity in normal and diseased tissues, but it possesses several challenges. As the amount of scRNAseq data present in the public domain has grown substantially in the last few years, a platform is needed to enable the meta-analysis. Moreover, multiple different scRNAseq protocols are in use, and a number of additional modalities co-analyzed with RNA is growing. Another challenging aspect of the analysis of scRNAseq data is the correct cell type identification which is especially important for studying the immune system consisting of a wide range of cell types and states that might have very different functions. Here we present a custom pipeline that enables unified and standardized processing of single-cell transcriptomics data from various sources, utilizing different protocols (10x, Smart-Seq) and accompanied by additional modalities (VDJ-Seq, CITE-Seq, Perturb-Seq). The processing of every single study includes raw data mapping, standardized and strict quality control, data normalization, and integration. These steps are followed by cell clustering with subsequent cluster identification and description. The cell type annotation is then synchronized across all studies in the compendium. This approach makes the cell type annotation more accurate and enables a compendium-wide meta-analysis. Finally, the pipeline outputs are enriched with sample-level information (e.g., patient-level data), and data are integrated into a user-friendly analysis software - GENEVESTIGATOR, a high-performance visualization tool for gene expression data. Taken together, we have built a manually curated and globally normalized scRNAseq compendium mainly consisting of immune cells obtained from studies focused on immuno-oncology, autoimmune diseases, and other therapeutic areas. This deeply harmonized compendium represents an important asset for downstream ML and AI applications in pre-clinical biomarker discovery and validation.

Secure and scalable Health Data: Making real world data actionable

Andreas Kremer, Information Technology for Translational Medicine (ITTM) S.A., Esch-sur-Alzette, Luxembourg

Abstract:

Healthcare data is typically heterogeneous in nature and brings operational, technical, and methodological challenges. Wearables, sensors, smartphone apps, IT-based medical data management platforms profoundly change current healthcare models, with huge impact on all stakeholders and potential to increase patient's benefits and participation. Real-world data (RWD) also offers the possibility to derive novel insights on the use and performance of medicines in everyday clinical use, complementing rather than competing with evidence from randomized control trials. There is increasing interest in the use of real-world data (RWD) to support clinical studies and regulatory decision making across the product life cycle. Key sources of RWD are electronic health records, claims data, prescription data, and patient registries. Increasingly incorporated into the definition is data from wearables, m-health apps, and environmental data including data on social status, education, and other lifestyle factors. There is a growing number of databases in healthcare organizations which contain this type of patient data. Still, to use this data optimally, we need to facilitate the collection of high-quality data and to foster standardized models and queries.

German Network for Bioinformatic Infrastructure

Responsible for the de.NBI Industrial Forum: Alfred Pühler, Nils-Christian Lübke

Head of de.NBI Administration Office
Head of Node ELIXIR Germany
Prof. Dr. Andreas Tauch
Research Center Jülich
Institut für Bio - und Geowissenschaften - IBG -5
Branch Office Center for Biotechnology / CeBiTec
Bielefeld University
Universitätsstr. 27
33615 Bielefeld

Deputy Head of Node ELIXIR Germany
Prof. Dr. Alfred Pühler
Bielefeld University
Center for Biotechnology / CeBiTec
Bielefeld University
Universitätsstr. 27
33615 Bielefeld

de.NBI Industrial Forum Coordination
ELIXIR Germany Industry Officer
Nils-Christian Lübke
Research Center Jülich
Institut für Bio - und Geowissenschaften - IBG -5
Branch Office Center for Biotechnology / CeBiTec
Bielefeld University
Universitätsstr. 27
33615 Bielefeld

contact@denbi.de   www.denbi.de