Building data bridges between biological and medical infrastructures in Europe

Results in Brief

A common infrastructure for all biomedical research

For every scientific field that could benefit from data integration to bridge basic and application-driven research, one of the most difficult points on the development agenda is that of interoperability. An EU-funded project has made important steps forward in this regard, with tools specifically aimed at biomedical researchers.

Bringing together 12 European biomedical sciences research infrastructures, BIOMEDBRIDGES (Building data bridges between biological and medical infrastructures in Europe) aims to remove the stumbling blocks to data sharing and interoperability within and across life science domains. It bridges data across different spatial scales, species, technologies and research communities to enable new ways of analysing problems and eventually answer new, more complex scientific questions.

To do so, the project relies on two groups of tools and specific use cases in order to demonstrate their power: on one hand, tools aimed at biomedical researchers across a number of disciplines, and on the other hand tools more specific to certain user communities or research questions.
An example of the first type of tool is the BIOMEDBRIDGES Meta Service Registry. It brings a variety of biomedical software and tools together to one single point of access and makes it easy for researchers to find, compare and use biomedical software and tools. It can help them address a scientific question or research support task such as “What are all of the Gene Ontology tools?” or “Which of these is most highly cited?” By returning relevant, structured results, the registry complements search engines like Google: the user can specify exactly what they need, using various search and filter options, and get a tailored list of suitable resources. From sequencing to structures, imaging to indexing, the registry’s domain scope is very broad, ensuring coverage of most parts of the BIOMEDBRIDGES partners’ activities, explains Friederike Schmidt-Tremmel, project manager at ELIXIR and coordinator of BIOMEDBRIDGES.

Another example is the ‘Legal and ethical requirements assessment tool’ (LAT), which guides researchers on how to work with sensitive data without necessarily consulting experts. It clarifies if and how a specific type and form of data can be shared and when additional actions or expert advice are needed — saving precious time in the process.

The DIAB Ontology is an example of the second group of tools, bridging the gap between mouse models and human studies in Type 2 diabetes and obesity. ‘Researchers working on human patients and those working on mouse models belong to different, mostly separate communities, each of which has its own ontology. There are over 100 human “Genome-wide association studies” (GWAS) annotated with “Diabetes” and over 750 mouse models (phenotypes) annotated with “increased circulating glucose level”,’ Schmidt-Tremmel says, continuing: ‘The DIAB ontology tool “crosses the species bridge” between mouse models and humans by establishing a diabetes-specific ontology for both and opening up extensive mouse phenotype data to clinical researchers. Clinicians can now compare human genomes with those of a well-established experimental model — the mouse — showing the same condition, and look deeper into the pathways behind glucose metabolism.’

With these computational bridges, the project aims to accelerate the research process from basic science to market applications, for example drug discovery. ‘BIOMEDBRIDGES extended the UniChem tool, originally developed as part of the EU-OPENSCREEN infrastructure to interlink between many different chemistry databases on small molecules. BIOMEDBRIDGES has developed the connectivity search function that allows users to find not only the “same” chemical compounds across different resources, but also “similar” compounds that differ in some of the characteristics. In this way, the UniChem tool finds and links 60 million related molecules from 21 data sources worldwide, including information on whether a specific compound has been patented or what research has been done on it,’ Schmidt-Tremmel explains. ‘This functionality can boost research into the mechanism of action of an existing drug and its possible off-label uses, which is
particularly important in the development of new pharmaceuticals, where early candidate triage (filtering out compounds that are most worthwhile to pursue) can save significant amounts of time and money.’

A total of five use cases were established, looking at data integration from the perspective of their user community or a concrete research topic: PhenoBridge, which crosses the species bridge between mouse and human; the interoperability of large-scale image data sets from different biological scales; personalised medicine (integrating complex data sets to understand disease pathogenesis and improve biomarker and treatment selection); cells to molecules (integrating structural data); and finally the integration of disease-related data and terminology from samples of different types.

A layered approach

BIOMEDBRIDGES has taken a layered approach to the integration of available data: interoperability is achieved via harmonising resources across research infrastructures. This initially involves using established technologies, but over the long run the project aims for more sophisticated semantic interoperability and user interfaces that will enable researchers to find — in one single step — information most relevant to a scientific question or a specific disease across millions of data entries.

‘This stepwise, layered approach ensures that all research infrastructures and data resources can systematically be brought to a higher level of integration,’ Schmidt-Tremmel says. ‘Almost as a side effect, this creates the necessary expertise of all involved to further advance data interoperability in future efforts. This continued collaboration also provides the opportunity to achieve real integration — possibly even a change of culture — across the life sciences with respect to data.’

The next step: CORBEL

Many of the tools developed by BIOMEDBRIDGES have already been released and are now available through the participating research infrastructures. They have been or will be embedded in different services provided by the contributing research infrastructures. For example, the knowledge gained from the development of the BIOMEDBRIDGES Meta Service Registry prototype has been absorbed into the ELIXIR Tools and Data Service Registry (bio.tools) and the LAT will become part of a more comprehensive service to support users and providers of sensitive data as part of ‘Ethical, legal and social implications’ (ELSI) common services to be developed within the new cluster project CORBEL.

Started in September 2015, CORBEL build on the results of BIOMEDBRIDGES to create a platform for harmonised user access to biological and medical technologies, biological samples and data services required by cutting-edge biomedical research.
Discover other articles in the same domain of application
Virtual tool for treating epilepsy

9 February 2023

Using AI to treat cancer

3 December 2021

Enhanced digital tool offers better support for epilepsy surgery decisions

3 March 2021

Last update: 11 July 2016

Permalink: https://cordis.europa.eu/article/id/157736-a-common-infrastructure-for-all-biomedical-research

European Union, 2024