PRESS RELEASE

Connecting the Dots between Cancer Data Networks: The SciLake Cancer Knowledge Graph

April 30th, 2024

Researchers from the Centre for Research and Technology in Greece and the Department of Molecular Medicine and Surgery, Karolinska Institutet in Sweden set up a collaboration to make publicly available resources in biology and cancer more accessible to the research community. The teams are developing a first-of-its-kind cancer knowledge graph (KG), as part of the SciLake Project involving partners across Europe.

The focus of this pilot project is to harness the power of SciLake services to aid in identifying biomarkers essential for personalised treatment and care, a crucial step towards the realisation of precision medicine. The case study of interest is Chronic Lymphocytic Leukemia (CLL), the most common adult leukaemia, characterised by a highly heterogeneous clinical course. The goal of this pilot is to gain a better understanding of this heterogeneity by leveraging information from the CLL knowledge graph (KG), a cancer-specific KG that can effectively model the respective knowledge space using defined data sources. CLL KG will combine existing biomedical KGs like the Clinical Knowledge Graph with text mining and entity recognition services to reveal connections between clinical covariates, genes, proteins, metabolites, mutations, drugs, scientific literature, biomedical datasets, software and research consortia.

The potential applications of this cancer KG are numerous. On the one hand, it will enable researchers to query existing information to deepen their understanding of their findings. On the other hand, it will enable more advanced users to use the KG to uncover latent knowledge using more advanced graph mining algorithms and machine

learning approaches. In the context of precision medicine, the CCL-KG will enable new discoveries of patient subtypes and why, for example, some groups respond better to certain treatments than others. In addition, researchers can use the KG as a benchmark to gauge the novelty of their discoveries in the field.

The teams have been diligently searching for existing knowledge graphs in the context of cancer. Recognising the ongoing need for an updated, cancer-specific knowledge graph, they have invested significant time in sourcing more general knowledge graphs from biology and precision medicine that will serve as a starting point. The challenge ahead lies in the effective integration and enrichment of these graphs to create a comprehensive cancer graph.

However, the potential benefits far outweigh the challenges. A successful cancer KG will be a significant asset for the wider cancer community. This tool, leveraging graph-based algorithms, user-friendly analytics, diverse clinical-biological databases, and drug-specific information, will empower scientists in their research.

About Karolinska Institutet

Karolinska Institutet is one of the world's leading medical universities. Its vision is to advance knowledge about life and strive towards better health for all. Karolinska Institutet accounts for the single largest share of all academic medical research conducted in Sweden and offers the country's broadest range of education in medicine and health sciences. The Nobel Assembly at Karolinska Institutet selects the Nobel laureates in Physiology or Medicine.

About CERTH

The Centre for Research & Technology HELLAS (CERTH) is a leading research centre in Greece conducting specialised basic and applied research and offering high quality services in Life Sciences. The Institute of Applied Biosciences at CERTH (INAB|CERTH) is dedicated to the promotion and execution of applied biosciences. Research at INAB|CERTH revolves around molecular medicine, biomedical informatics, and the development and evaluation of eHealth, mainly in chronic diseases.

About SciLake

SciLake is a project funded by the European Union's Horizon Europe program. The project aims to create a seamless integration between domain knowledge and open Scientific Knowledge Graphs while also developing useful added-value services for specific research areas. The ultimate goal is to empower researchers and foster a more interconnected and efficient scientific community. SciLake brings together a competent consortium of 13 partners from 9 different countries. The consortium consists of partners with expertise in knowledge management and discovery, as well as experts from Neuroscience, Cancer, Transportation, and Energy research, who are involved in piloting activities.

Contacts:

Dania Machlab	dania.machlab@ki.se
Leily Rabbani	leily.rabbani@ki.se
George Gavriilidis	ggeorav@certh.gr
Konstantinos Kardamiliotis	k.kardamiliotis@certh.gr
Fotis Psomopoulos	fpsom@certh.gr
Vasileios Vasileiou	vasileioubill95@certh.gr

###