Building a secure data processing environment for personalized health research in Switzerland

The Swiss Personalized Health Network (SPHN) is a national initiative designed to promote the development of personalized health in Switzerland. A main goal is to bring Switzerland at the forefront of personalized health research by establishing nationwide interoperability of biomedical information. In particular, the BioMedIT project will establish a coordinated network of secure infrastructures to support computational biomedical research and clinical bioinformatics. BioMedIT builds on established centers of expertise at partner universities ("nodes"), coordinated by the SIB Swiss Institute of Bioinformatics. BioMedIT nodes provide scientific and technical expertise in computational life science research, pre-installed software and workflows, large-scale storage and reference data for "big-data" applications, and high-performance computing capacity to enable state-of-the-art life science research.

The Swiss Data Science Center (SDSC) is part of the Initiative for Data Science in Switzerland launched by the ETH Domain. The SDSC addresses the gap between those who create data, those who develop data analytics and systems, and those who could potentially extract value from it, while fostering innovation in data science, multidisciplinary research and open science. As part of its mandate, the SDSC is developing RENGA, an open-source distributed software platform designed to facilitate the exchange of potentially sensitive data and knowledge between all the actors involved in data science collaborations, while enforcing their respective data management plans.

The SIB and the SDSC are teaming up on developing software platforms, services and best practices for privacy-conscious data-driven science in the context of SPHN. Through this collaboration, the joint team will assess the domain specific requirements for personalized health research in Switzerland. The joint team will then develop and operate a domain specific extension of RENGA according to these requirements.

One key aim of this collaboration is to leverage RENGA to enable data traceability and transparency in the form of provenance information. This enables scientists to validate the veracity and reproducibility of computational results. Provenance will also allow them to replicate the computation and confidently reuse the results in derivative work with all the proper accreditation and attribution. In the health data domain, traceability of data is required to comply with legal requirements for data protection.

From a technical perspective, future versions of the RENGA platform will enable a federated mode of operation. Regional BioMedIT nodes are operating IT infrastructure such as large-scale storage, high-performance computing, software and workflows supporting scientific projects and analytical platforms. Individually managed instances of RENGA at these nodes will allow joining resources to an interconnected data processing ecosystem for exploratory data analysis, while nodes remain in full control of the permission settings on their respective resources.

The aim of this strategic partnership is to develop a one-stop-shop to researchers and clinicians for accessing and analyzing health data while conforming to applicable data governance regulations, eventually enabling collaborative and reproducible exploratory research in a privacy-conscious environment.

About the SIB Swiss Institute of Bioinformatics :

The SIB Swiss Institute of Bioinformatics is an academic not-for-profit organization whose mission is to lead and coordinate the field of bioinformatics in Switzerland. Its data science experts join forces to advance biological and medical research, and enhance health. SIB (i) provides the national and international life science community with a state-of-the-art bioinformatics infrastructure, including services, resources, expertise; and (ii) federates world-class researchers and delivers training in bioinformatics. The institute includes some 70 world-class research and service groups including 800 scientists in the fields of genomics, proteomics, evolution and phylogeny, systems biology, structural biology, text mining and machine learning as well as personalized health.

About the SDSC Swiss Data Science Center :

The Swiss Data Science Center is a joint venture between EPFL and ETH Zurich. Its mission is to accelerate the adoption of data science and machine learning techniques within academic disciplines of the ETH Domain, the Swiss academic community at large, and the industrial sector. In particular, it addresses the gap between those who create data, those who develop data analytics and systems, and those who could potentially extract value from it. The center is composed of a large multi-disciplinary team of data and computer scientists, and experts in select domains, with offices in Lausanne and Zurich.
www.datascience.ch

About the SPHN Swiss Personalized Health Network :

The Swiss Personalized Health Network (SPHN) is a national initiative designed to promote the development of personalized medicine and personalized health in Switzerland. SPHN will lay the foundations that are needed to facilitate research projects in this area such as a system for a nationwide exchange of health-related data. The ultimate goal is to promote health and well-being, to prevent, diagnose and treat unfavourable health conditions more precisely, thus reducing the risk of developing such conditions and permitting more effective treatments of disease states with fewer adverse effects.
www.sphn.ch

Contact information :

SDSC	SIB	SPHN
Olivier Verscheure	**Torsten Schwede**	**Dr.Adrien Lawrence**
Executive Director	Director of SPHN Data Coordination	Managing Director SPHN
olivier.verscheure@datascience.ch	Center & BioMedIT project	info@sphn.ch
+41 21 693 26 50	torsten.schwede@sib.swiss	
	+41 61 207 15 81	