

Mandate

SPHN IT Architecture Working Group

31 August 2021, approved by HIT-STAG

Starting position

The Swiss Personalized Health Network (SPHN) promotes the development, implementation, and validation of coordinated infrastructures to make health data interoperable and shareable for research in Switzerland. As part of the SPHN initiative, various efforts are currently underway to harmonize and define data standards in order to ensure the interoperability of health-related data, with a specific focus on data from the five Swiss University Hospitals (UH). In the first funding phase, SPHN partners started the development of several infrastructure elements, among them:

- Initial hospital specific infrastructure to provide clinical data;
- Semantic definition of (core) data elements;
- Introduction of RDF as data exchange format with related SPHN schema definition;
- Setup of the BioMedIT network as secure IT environment for the processing of sensitive research data;
- Onboarding of various Swiss data providers to the BioMedIT network and development of tools and services for the secure transfer and management of health data;
- Setup of an initial Federated Query System containing over 56.3 million data entries of 5 UH;
- Legal agreement templates for the contractual framework necessary to conduct multicenter research projects involving the exchange of health-related data.

From the currently running SPHN Driver Projects and the work in relation to the [SPHN semantic interoperability framework](#), SPHN has identified a series of critical and challenging issues. Among them is the need for a harmonized IT architecture in order to streamline processes concerning data delivery to research projects. The aim is the reduction of redundant work steps regarding, for example, standardization efforts, data preparation, quality assurance, and data representation according to the required specifications. In particular, the following challenges and gaps have been identified and need to be tackled:

- Centrally guided architecture in form of an IT lead architect has been missing so far, covering all SPHN infrastructure aspects;
- Structured data is still rare at hospitals;
- RDF as data exchange format is cost intensive to build at hospitals;
- RDF as data exchange format is difficult to use for researchers, who are not data scientists;
- RDF as exchange format currently is based on an SPHN specific schema definition; intensified alignment with operative routine data exchange formats at hospitals should be considered to benefit from synergistic effects (e.g., HL7; upcoming: FHIR);
- Promised benefit of currently established Federated Query System needs to be confirmed.
- Non-harmonized IT SPHN related architecture at the hospital side leads to multiple redundant infrastructure components with various results and qualities; e.g. RDF extractors.

A common architectural framework would address many of the existing problems and inefficiencies and would have an impact on the sustainability of the established infrastructure, particularly in view of the post-SPHN funding period. Having this common architectural framework in place would allow the hospitals to focus their work on data quality and new data pipelines rather than having to maintain redundant solutions.

Vision and mission

The vision is to have streamlined processes and a cost-efficient sustainable way to convey structured data from UH to a central research environment in accordance with the SPHN Interoperability Strategy, making the necessary steps for the university hospitals in their role as data providers smooth and efficient.

The envisioned architecture should not be geared exclusively to the needs of SPHN and related research projects, but – due to synergies and cost efficiency – should also consider internal needs of the UH (e.g. own research) and local research initiatives (such as The Loop in Zurich), as well as additional cases where hospitals are obliged to provide data to external parties (e.g. the feeding of registers, public health data collections, data deliveries to authorities, etc.). With regard to the architecture to be defined, the “once-only” principle should be the most important premise, meaning that harmonization efforts allow:

- Foster common infrastructure elements as far as possible for all stakeholders instead of multiple stakeholder-specific solutions for the same purpose;
- Limit effort for Data standardization at hospitals; ideally it would be executed only once and already at the point of data entry or be managed centrally, in order to prevent standardization and mapping efforts at different sites in one hospital.

Key requirement of the envisioned architecture is the sustainability aspect with regard to future data deliveries/provisions to research projects in alignment with operative routine data exchange formats at hospitals, also beyond the funding phase of SPHN. However, also the current requirements need to be addressed in the envisioned architecture, namely the presence of a Federated Query System over the 5 UH, the option for privacy preserving data exploration (for UH internal and external researchers), the preparation and delivery of data outside the UH for complex SPHN driver/lighthouse projects (in the realm of the NDS to come).

Tasks and deliverables of the Working Group

- Constantly manage **requirements in a requirements repository** and consider them in an architecture plan.
- Provide an **architecture and maintenance plan for a hospital data delivery pipeline** which provides high quality data and is sustainable and efficient for hospitals to develop and operate. It should take clinical routine standards into account, to avoid additional overhead.
- Provide an **architecture and maintenance plan for a sustainable and secure research infrastructure**, which is based **on top** of the above mentioned hospital data delivery pipeline, including appropriate research representation of clinical data (e.g. RDF-format, Graph-representation), Federated Query System, Metadata Catalogue for clinical data, Cohort Manager, Data Transfer Pipelines, etc., and which provides high quality data to SPHN research projects (fit-for-purpose).
- Provide a **roadmap** for implementation of the overall architecture. Together with the project manager, the HIT-STAG and additional stakeholders, a detailed timeline will be defined including hospital resources as well as other required resources e.g. on the DCC side, the research side, on BioMedIT, etc.

- **Initiate single projects** for implementing the architecture, i.e. build infrastructure and related maintenance organization. Projects have to be executed by a dedicated project manager. All projects in the SPHN portfolio should be coordinated by a SPHN Program Manager.
- Ensure **Implementation Governance of single projects** (e.g. infrastructure development projects, proof-of-concept projects) by revising project specific solution specifications and integrate project deliverables in an overall architecture plan.

Initial Architecture Principles

The concept of the overall architecture should specify aspects such as level of data standardization and de-identification, legal responsibility, functionalities and use cases/clients, processes and tools for a sustainable data-driven architecture governance.

Moreover, the concept should cover solutions for the identified gaps and challenges and consider the requirements of all stakeholders, especially hospital data providers, BioMedIT, researchers, and patients. The planned architecture must efficiently support delivery milestones defined in related contracts, such as the collaboration agreements between SPHN and the UH, HospFAIR program agreements, etc.

Methodology

- The architecture will be elaborated in an iterative approach.
- Project deliverables, representing SPHN central infrastructure elements, require appropriate specification and test documentation, as well as user and operation manuals (e.g. user requirements, functional and technical specifications).

Composition and organization of the WG

The HIT-STAG will nominate one working group member per UH. The DCC will provide project management support and program lead (up to 1 FTE).

Additional to the fixed working group members, there are several additional advisory members that can be called in WG meetings as needed, i.e.

- user/research representatives (e.g. clinical scientists, data scientists, medical informatics experts)
- subject matter experts (e.g. on technical exchange formats [HL7, FHIR, RDF], semantic experts, information security experts, metadata catalogue experts),
- BioMedIT node representatives
- external consultants (if deemed necessary).

Since the position of an SPHN Lead Enterprise Architect could not be filled, the lead of the WG and main the responsibility for the development of the deliverables will be given to volunteering UH representatives.

Once the Boards approved the Architecture concepts, the organization of the implementation workforce will be discussed and proposed to the HIT-STAG in the realm of the implementation roadmap.

The working group shall report to the HIT-STAG for deliverable discussions and approvals and in case of escalation.

Timelines

- Setup initial organization of architecture group (15.09.2021)
- Architecture principles aligned with the overall requirements (31.10.21)
- Refined high level concept definition (10.01.22)
- High level definition of work packages (WP) to specify the individual components (15.02.22)
- Define Roadmap for implementation projects with detail specification and realization (dependencies of WP, priorities) (31.03.22)
- After 31.03.2022, elaborate WPs in implementation projects:
 - Detail Specification of each WP
 - Implementation of the WP
 - Validation of implementation
 - Transfer to maintenance organization
 - Review and refine overall Architecture

The overall deliverables are planned to be delivered to the stakeholders in a stepwise approach until 31st March 2022. Approval will happen after the presentation of each deliverable to the stakeholder. Delayed approval process or rejection can/will have an impact on the planning.

Deliverable and deadline	Deadline for SPHN Boards	Comment / Funding
Presentation of the WG mandate by HIT-STAG: 31. AUG 2021	<ul style="list-style-type: none"> – Approval of the WG Mandate by HIT-STAG: 31. AUG 2021 – Nomination of UH representatives for WG until 6 SEP 2021. 	
1: Concept of the overall architecture <ul style="list-style-type: none"> – Hand-in to HIT-STAG: NOV 2021 	<ul style="list-style-type: none"> – Approval of the architecture principles by HIT-STAG: 10. NOV 2021 – Presentation to NSB (for information): 25. NOV 2021 – Approval of the Concept incl. WP definitions: FEB 2022 	Funding: <ul style="list-style-type: none"> – Reimbursement of the Lead Architect(s) from the volunteering UH by the DCC
2: Roadmap for implementation including required resources <ul style="list-style-type: none"> – First presentation of a draft to HIT-STAG: FEB 2022 – Hand-in to HIT-STAG: End of MAR 2022 	<ul style="list-style-type: none"> – Approval by HIT-STAG and NAB: APR 2022 – Approval by NSB: ~ MAI 2022 	Funding for the Roadmap development: <ul style="list-style-type: none"> – Reimbursement of the Lead Architect(s) from the volunteering UH by the DCC – Part of the HospFAIR program (potentially for first single projects) Funding sources for implementation in general (details tbd): <ul style="list-style-type: none"> – CA SPHN/UH – HospFAIR program – BioMedIT project funds