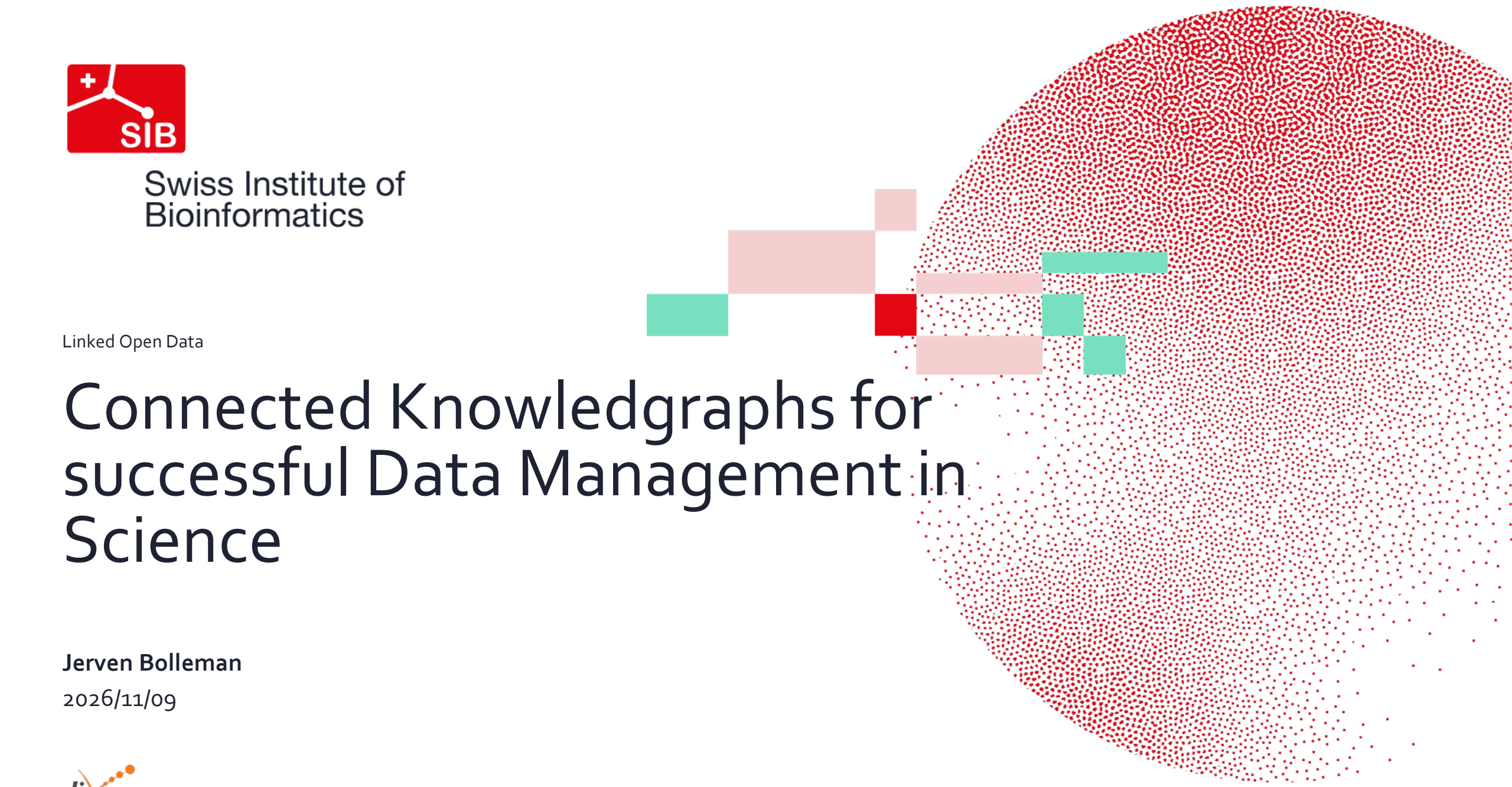


Swiss Institute of Bioinformatics

Linked Open Data

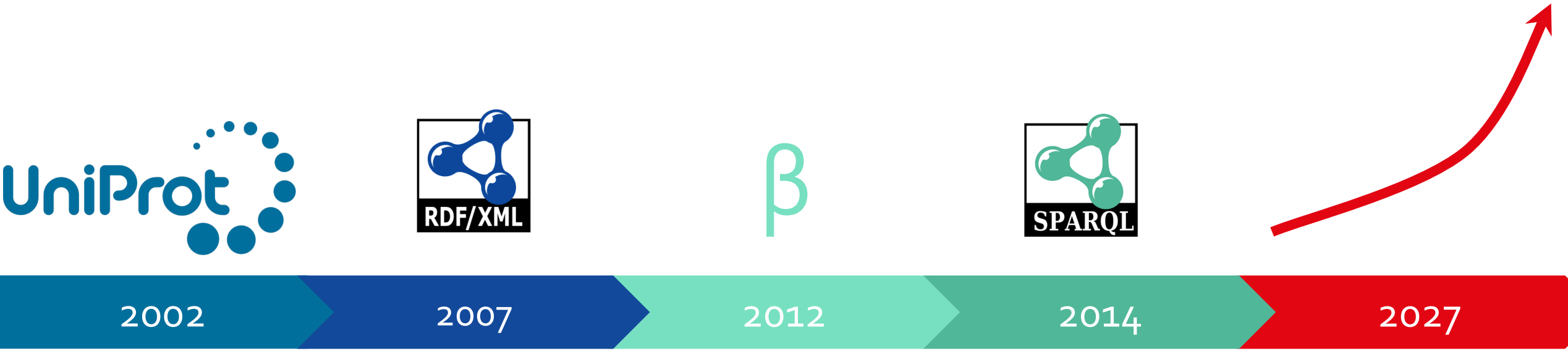
Connected Knowledgraphs for successful Data Management in Science

Jerven Bolleman

2026/11/09

elixir SWITZERLAND

https://sparql.uniprot.org/sparql a brief history



UniProt was founded by merging three existing databases. Combining their strengths we also went to free use for everyone instead of only free for academia.

RDF was being tested since 2004, but since 2007 it was in production and a key data product. About 5 GB in size to download.

In 2012 we started deploying a SPARQL endpoint. With over 5 billion triples it was already a challenging data set.

Took a while to select a SPARQL vendor. Virtuoso open source, by OpenLink, served us very well for the next 10 years.

After a decade in production our data keeps on growing, reaching 2 TB to download.

Number of users has grown as well.

Knowledge Graph not just a *Knowledge Base*

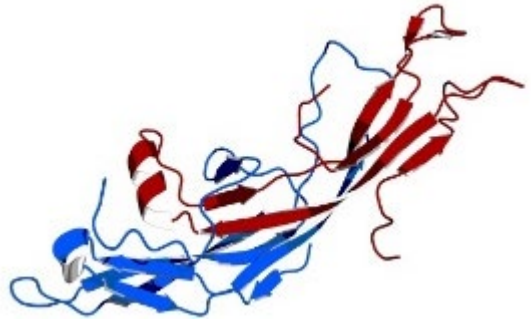
Why *Semantics* ?

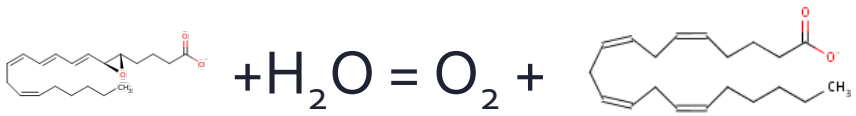
In Biology everything is connected

UniProt Rhea SwissLipids









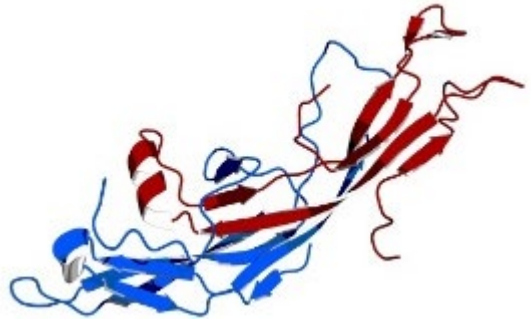
DNA → RNA → Proteins → Reactions → Metabolites

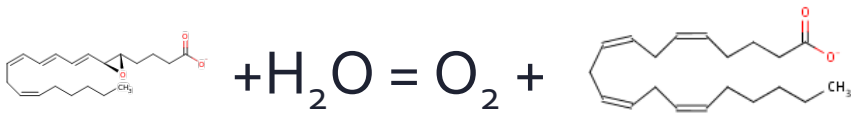
In Biology everything is connected

UniProt Rhea SwissLipids









DNA → RNA → Proteins → Reactions → Metabolites

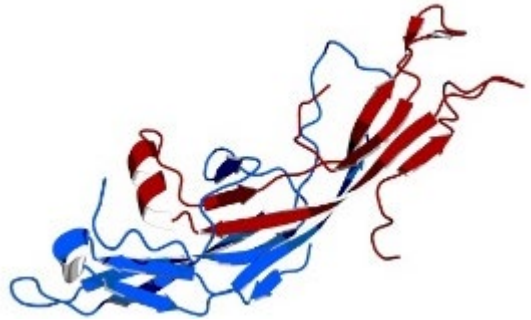
Disease

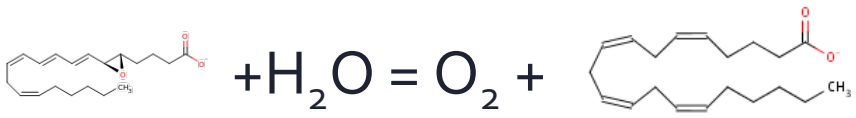
In Biology everything is connected

UniProt Rhea SwissLipids









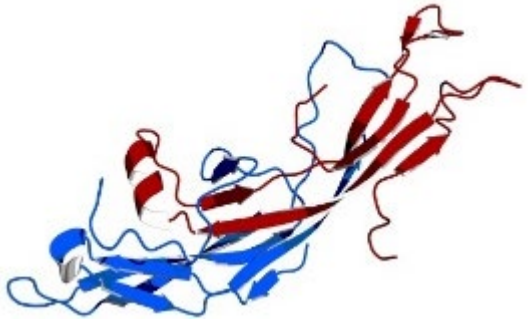
DNA → RNA → Proteins → Reactions → Metabolites

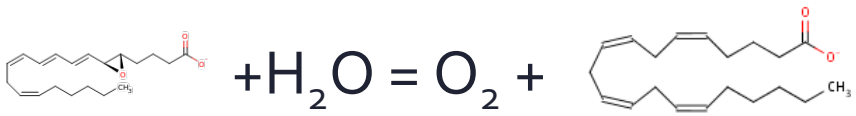
Disease

Medicine

In Biology everything is connected

UniProt Rhea SwissLipids

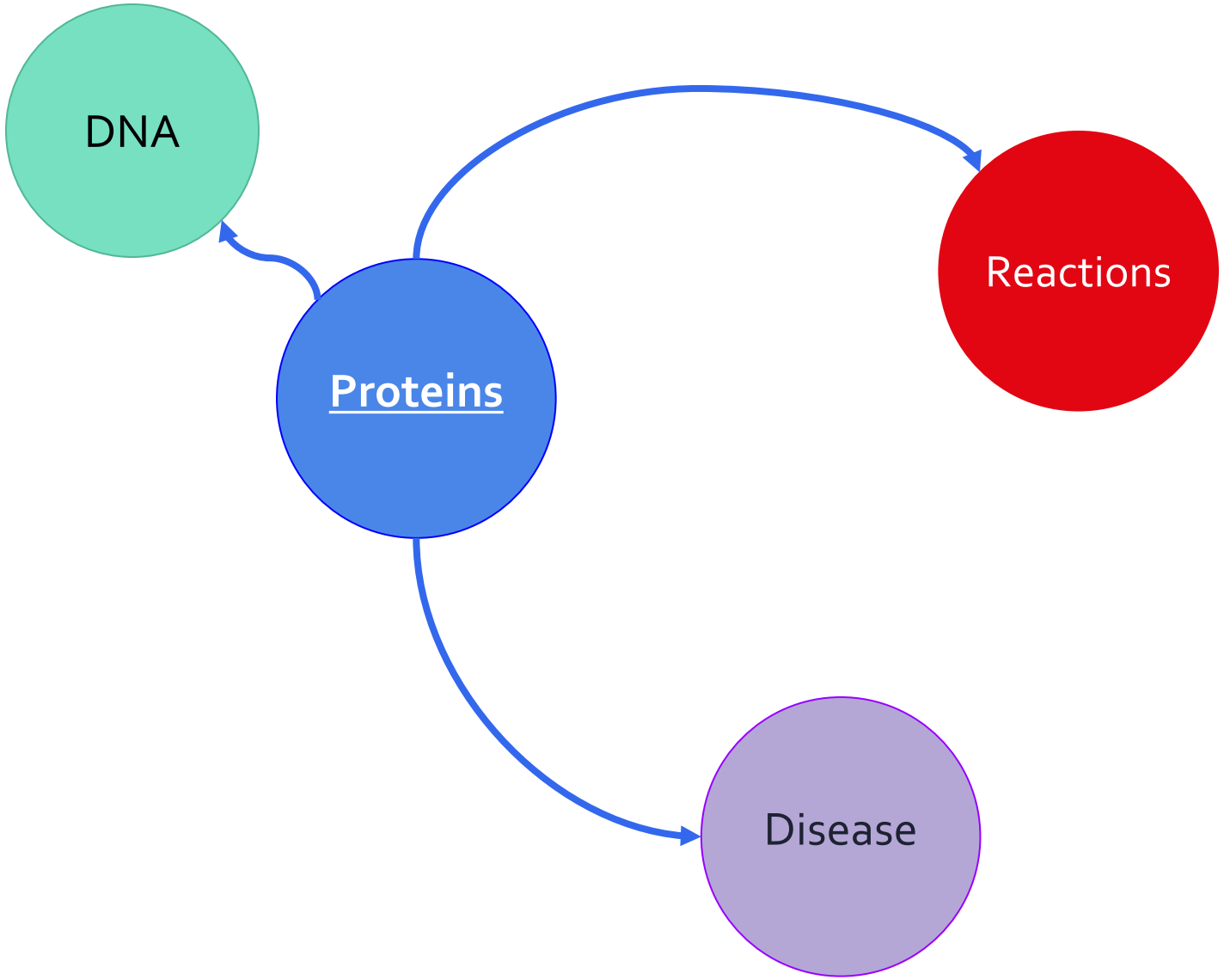




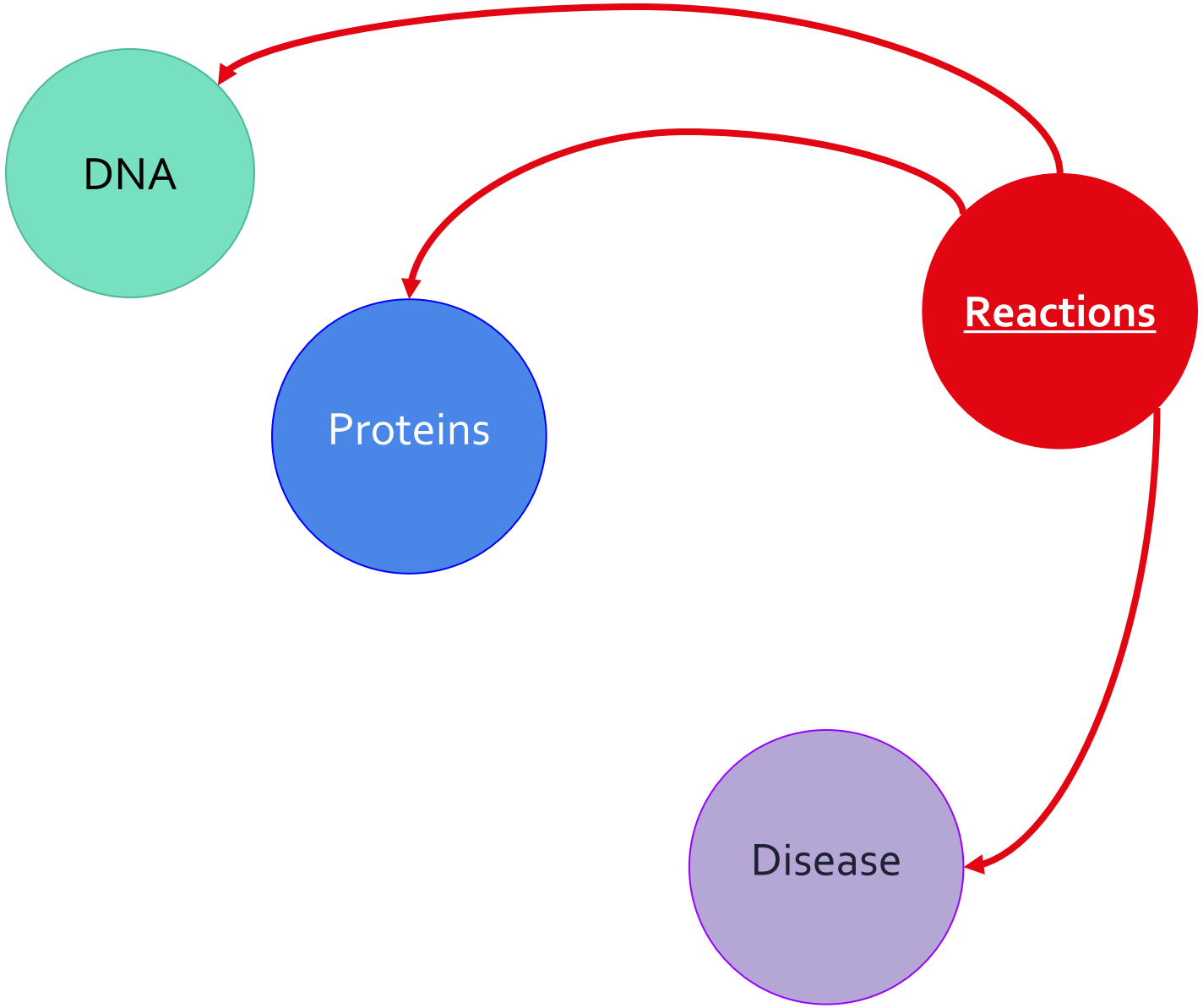
Proteins Reactions Metabolites

Disease

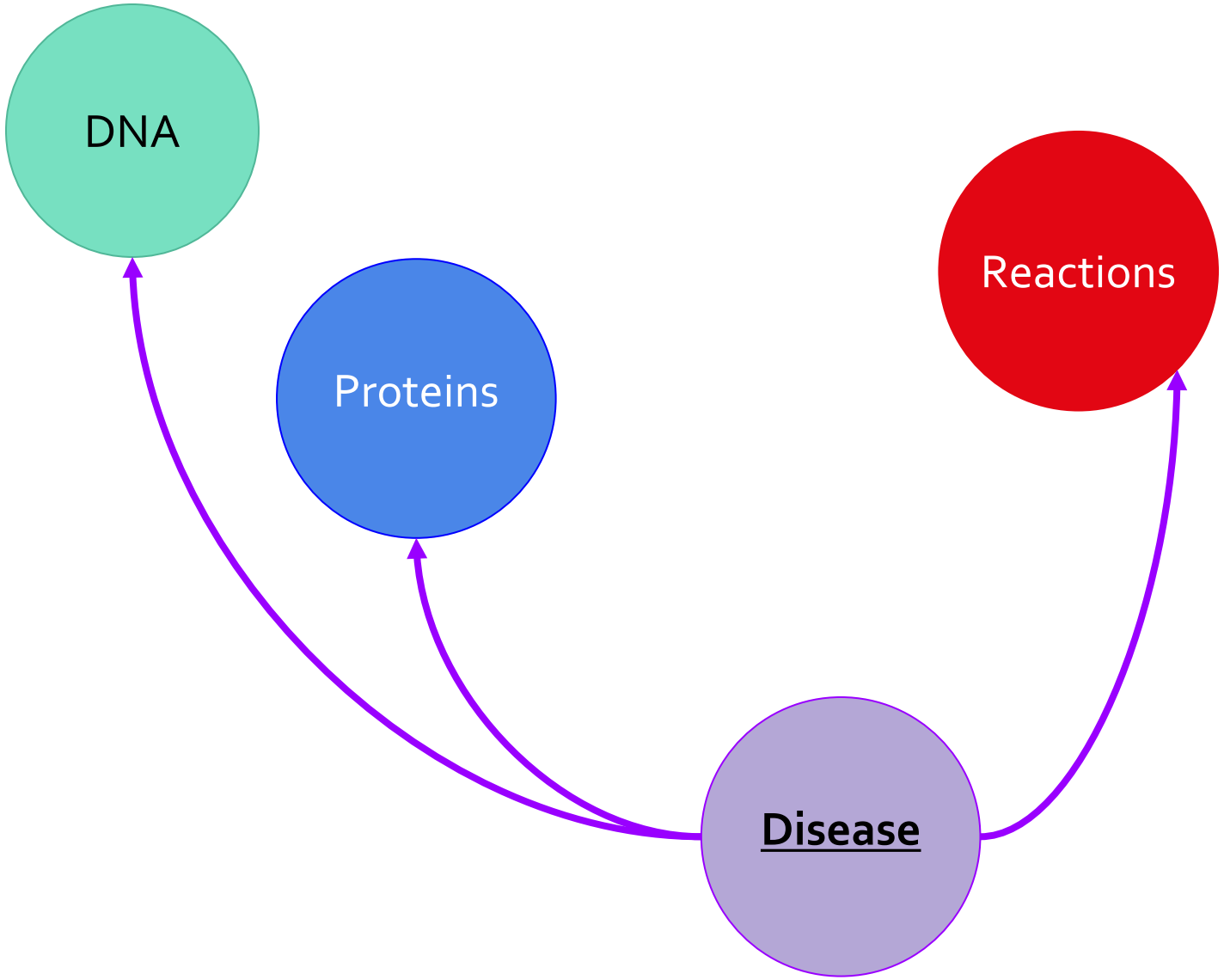
During research viewpoints change



During research viewpoints change



During research viewpoint change



Mission

Connecting science

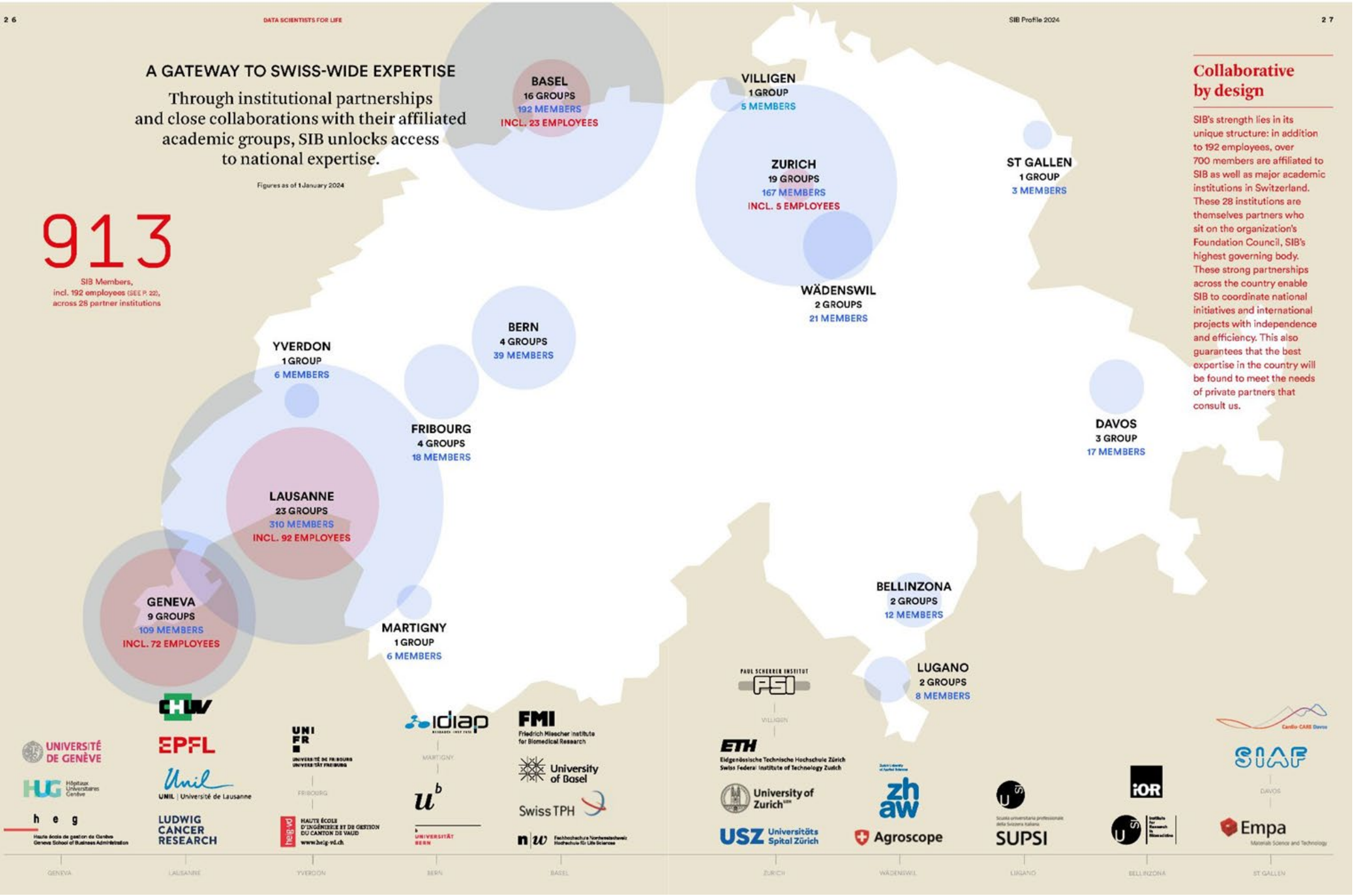
A GATEWAY TO SWISS-WIDE EXPERTISE

Through institutional partnerships and close collaborations with their affiliated academic groups, SIB unlocks access to national expertise.

Figures as of 1 January 2024

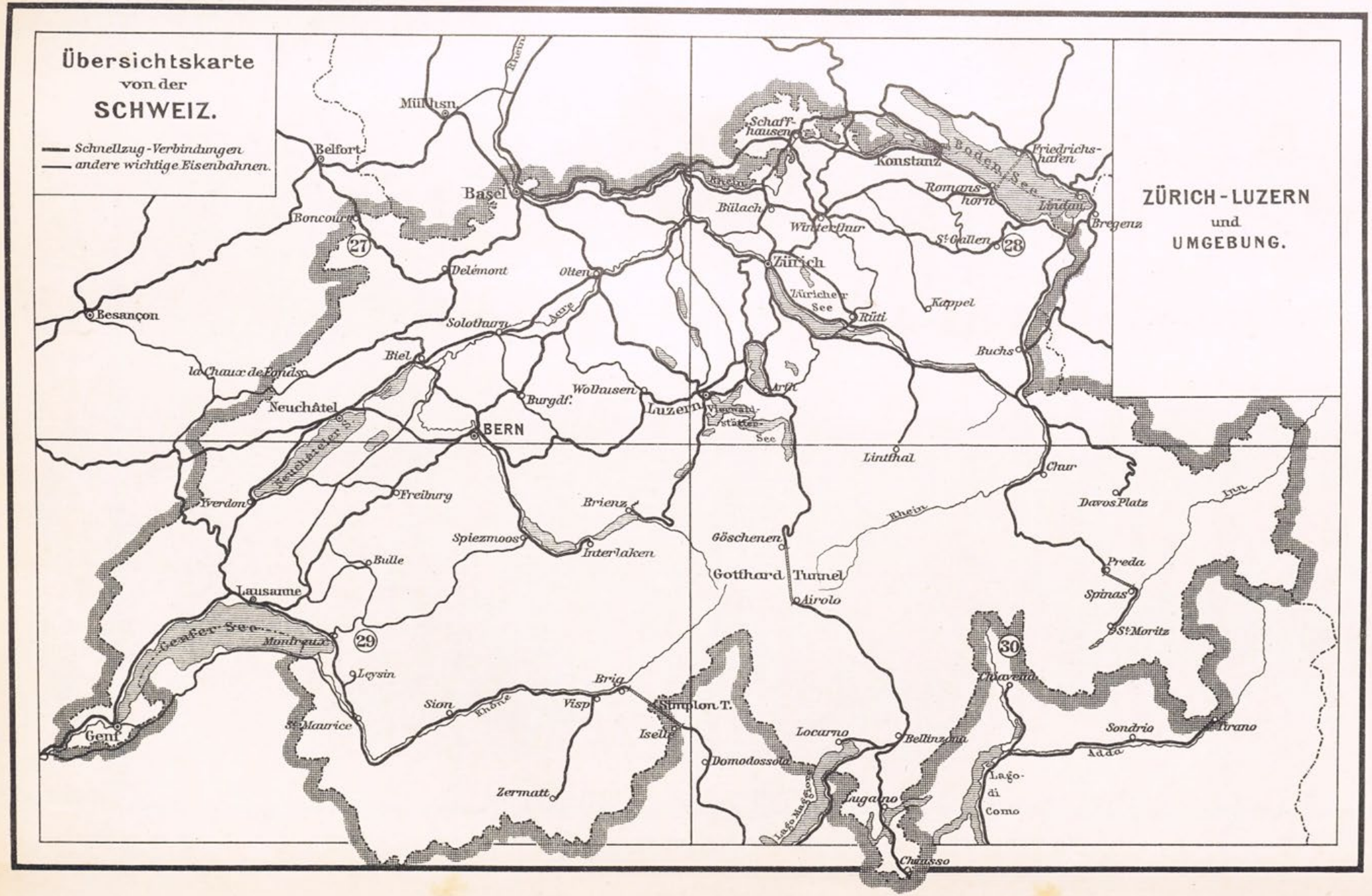
913

SIB Members, incl. 192 employees (SEE P. 22), across 28 partner institutions



Collaborative by design

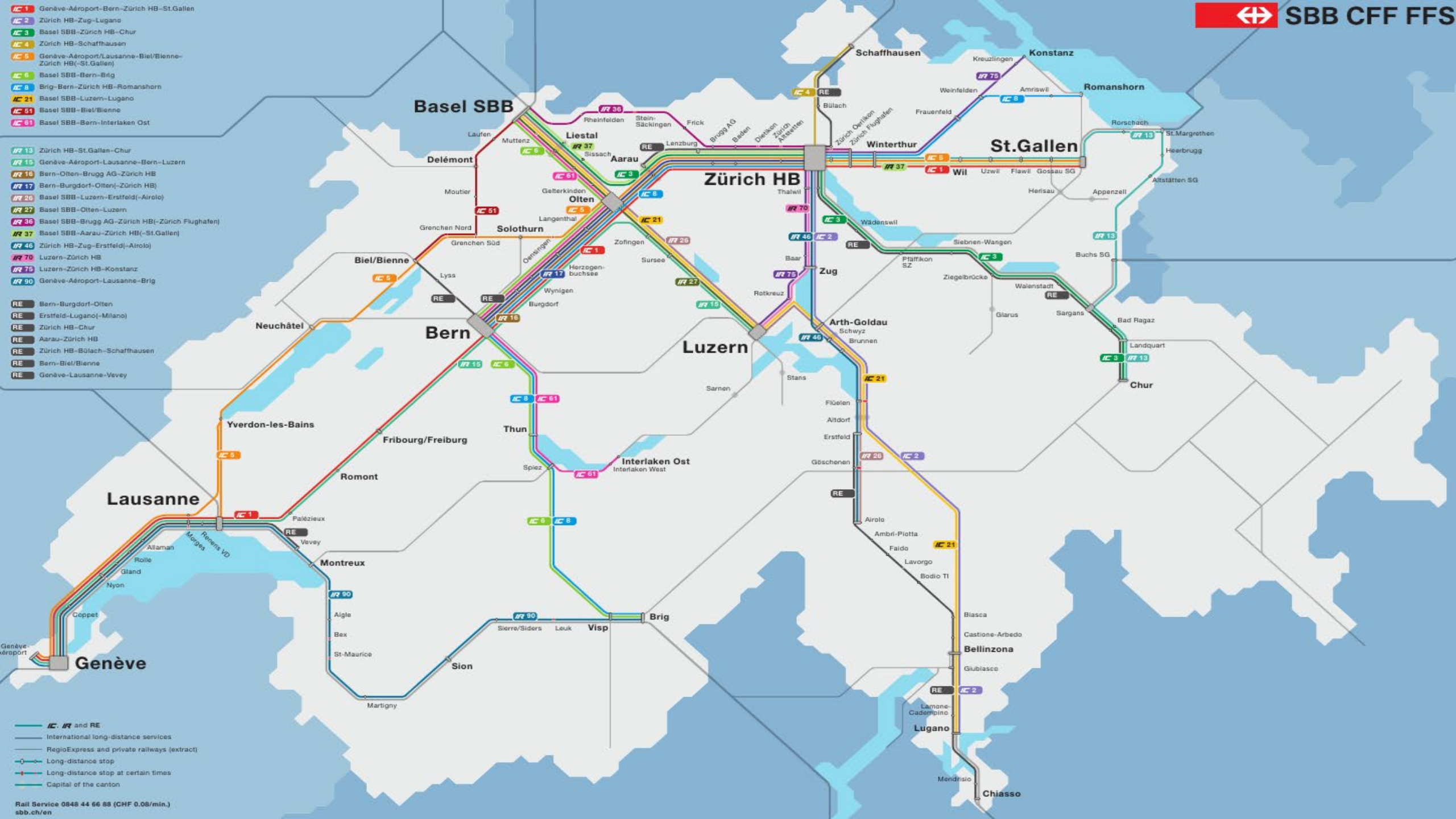
SIB's strength lies in its unique structure: in addition to 192 employees, over 700 members are affiliated to SIB as well as major academic institutions in Switzerland. These 28 institutions are themselves partners who sit on the organization's Foundation Council, SIB's highest governing body. These strong partnerships across the country enable SIB to coordinate national initiatives and international projects with independence and efficiency. This also guarantees that the best expertise in the country will be found to meet the needs of private partners that consult us.



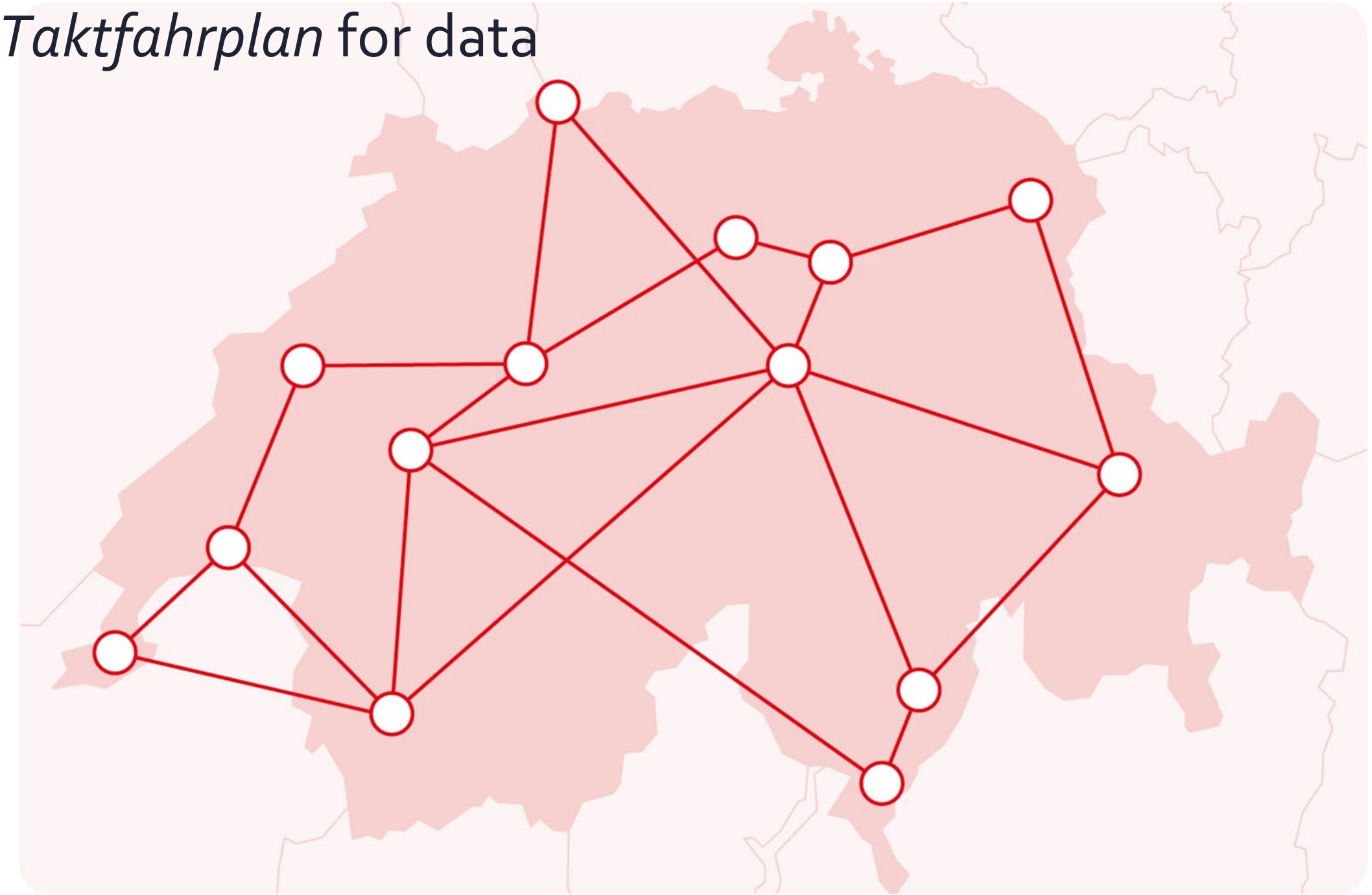
SCHWEIZ.

Section 1.


Schweiz. Einzelaufn. Nr.
LöS PS.
E 7/5
160
BERN
LÖTSCHBERG
SIMPLON
1256
Lok. Winterthur
BBC. Baden.
Prof. Wiesinger
Zürich.


SBB CFF FFS
IC 1 Genève-Aéroport–Bern–Zürich HB–St.Gallen
IC 2 Zürich HB–Zug–Lugano
IC 3 Basel SBB–Zürich HB–Chur
IC 4 Zürich HB–Schaffhausen
IC 5 Genève-Aéroport/Lausanne–Biel/Bienne–Zürich HB(–St.Gallen)
IC 6 Basel SBB–Bern–Brig
IC 8 Brig–Bern–Zürich HB–Romanshorn
IC 21 Basel SBB–Luzern–Lugano
IC 51 Basel SBB–Biel/Bienne
IC 61 Basel SBB–Bern–Interlaken Ost
IR 13 Zürich HB–St.Gallen–Chur
IR 15 Genève-Aéroport–Lausanne–Bern–Luzern
IR 16 Bern–Olten–Brugg AG–Zürich HB
IR 17 Bern–Burgdorf–Olten(–Zürich HB)
IR 26 Basel SBB–Luzern–Erstfeld(–Airolo)
IR 27 Basel SBB–Olten–Luzern
IR 36 Basel SBB–Brugg AG–Zürich HB(–Zürich Flughafen)
IR 37 Basel SBB–Aarau–Zürich HB(–St.Gallen)
IR 46 Zürich HB–Zug–Erstfeld(–Airolo)
IR 70 Luzern–Zürich HB
IR 75 Luzern–Zürich HB–Konstanz
IR 90 Genève-Aéroport–Lausanne–Brig
RE Bern–Burgdorf–Olten
RE Erstfeld–Lugano(–Milano)
RE Zürich HB–Chur
RE Aarau–Zürich HB
RE Zürich HB–Bülach–Schaffhausen
RE Bern–Biel/Bienne
RE Genève–Lausanne–Vevey
Basel SBB
Schaffhausen
Konstanz
Kreuzlingen
Romanshorn
Weinfelden
Amriswil
Bülach
Frauenfeld
Rheinfelden
Stein-Säckingen
Frick
Brugg AG
Baden
Dietikon
Zürich Altstetten
Zürich Oerlikon
Zürich Flughafen
Winterthur
St.Gallen
Rorschach
St.Margrethen
Heerbrugg
Altstätten SG
Laufen
Muttenz
Liestal
Sissach
Aarau
Lenzburg
Delémont
Moutier
Grenchen Nord
Grenchen Süd
Solothurn
Gelterkinden
Olten
Langenthal
Oensingen
Zofingen
Sursee
Zürich HB
Thalwil
Wil
Uzwil
Flawil
Gossau SG
Herisau
Appenzell
Wädenswil
Siebnen-Wangen
Pfäffikon SZ
Ziegelbrücke
Buchs SG
Walenstadt
Sargans
Bad Ragaz
Landquart
Chur
Glarus
Biel/Bienne
Lyss
Herzogenbuchsee
Wynigen
Burgdorf
Bern
Neuchâtel
Rotkreuz
Baar
Zug
Arth-Goldau
Schwyz
Brunnen
Luzern
Stans
Sarnen
Flüelen
Altdorf
Erstfeld
Göschenen
Airolo
Ambri-Piotta
Faido
Lavorgo
Bodio TI
Biasca
Castione-Arbedo
Bellinzona
Giubiasco
Lamone-Cadempino
Lugano
Mendrisio
Chiasso
Yverdon-les-Bains
Fribourg/Freiburg
Romont
Thun
Spiez
Interlaken Ost
Interlaken West
Lausanne
Palézieux
Vevey
Montreux
Aigle
Bex
St-Maurice
Martigny
Sion
Sierre/Siders
Leuk
Visp
Brig
Genève
Genève-Aéroport
Coppet
Nyon
Gland
Rolle
Allaman
Morges
Renens VD
IC, IR and RE
International long-distance services
RegioExpress and private railways (extract)
Long-distance stop
Long-distance stop at certain times
Capital of the canton
Rail Service 0848 44 66 88 (CHF 0.08/min.)
sbb.ch/en

Taktfahrplan for data



IRIs as *identifiers*
make connections possible

FAIR principle: findable 1

We must identify everything

What is 9993 ?

Coupling of the Penicillium Duponti Acid Protease to Ethylene-Maleic Acid (1 : 1) Linear Copolymer. Preparation and Properties of the Water-Soluble Derivative

PubMed



NCBI Taxonomy

DGCR2 DiGeorge syndrome critical region gene 2 [Homo sapiens (human)]

NCBI Gene

Is why we use IRIs

https://pubmed.ncbi.nlm.nih.gov/9993/

PubMed

http://purl.uniprot.org/taxonomy/9993



NCBI Taxonomy

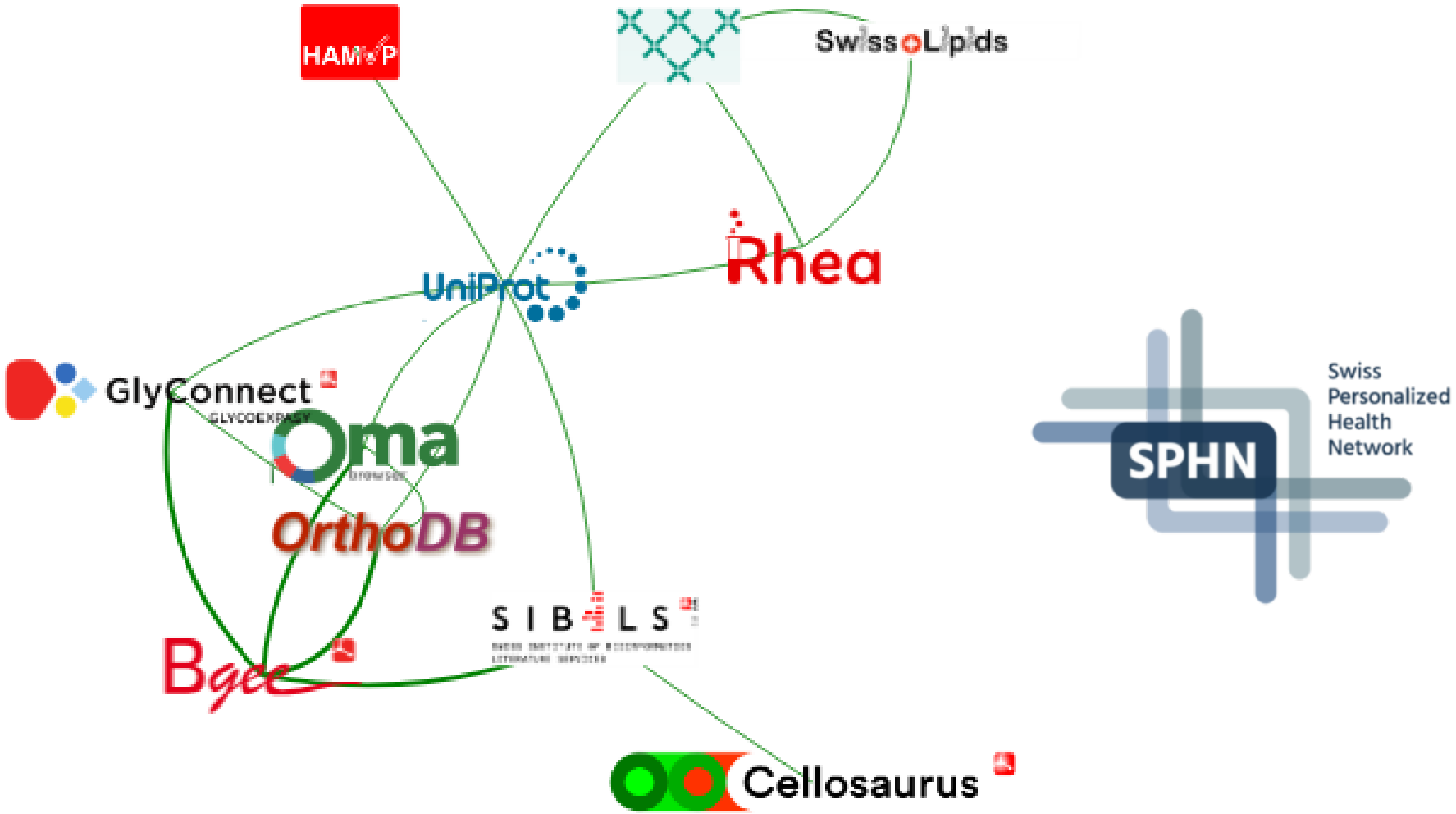
http://identifiers.org/ncbigene/9993

NCBI Gene

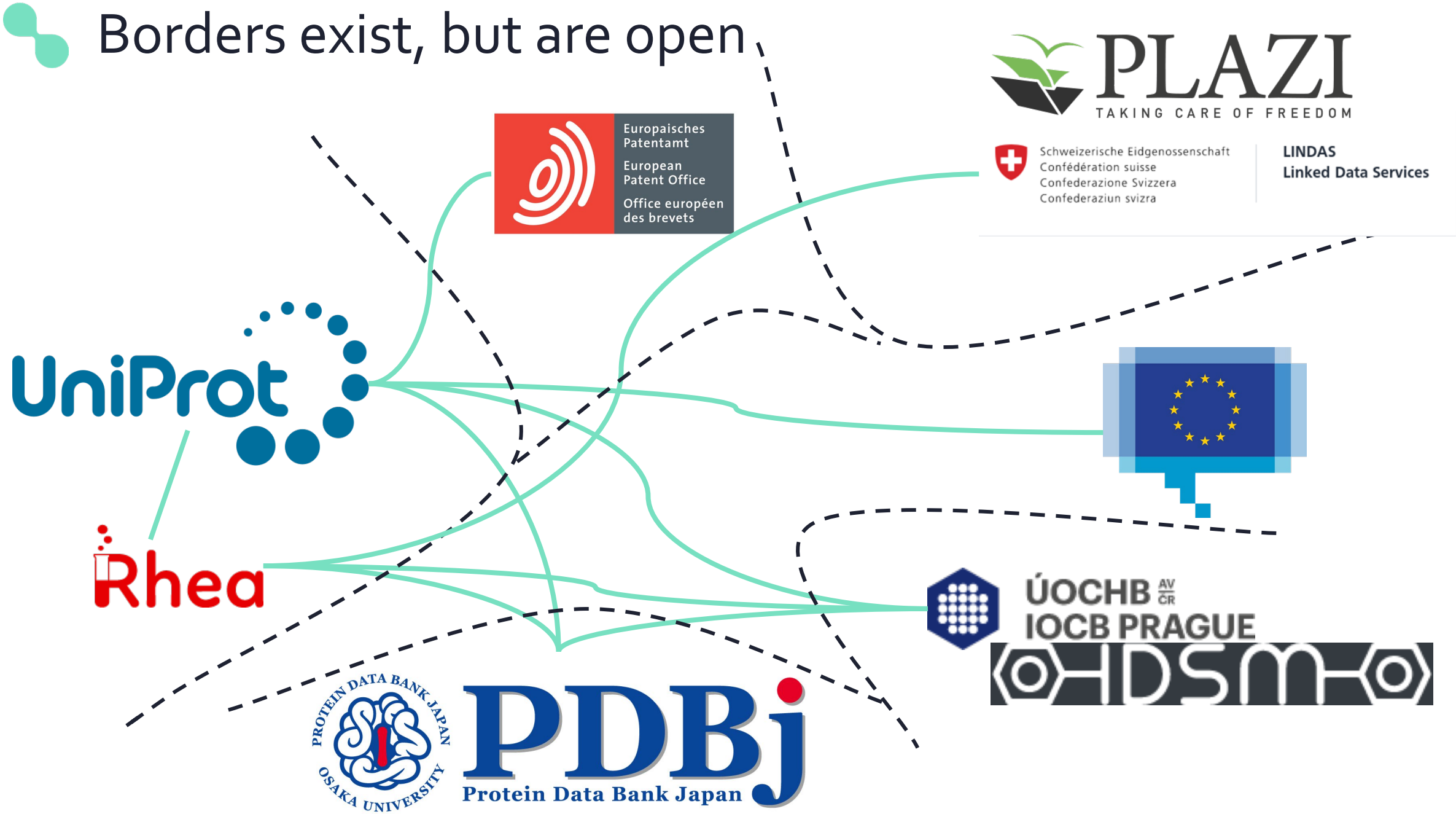
AS/NZS 3112

NEMA 5-15




HAMAP
SwissLipids
Rhea
UniProt
GlyConnect
Oma
OrthoDB
SIB
Bgee
Cellosaurus
SPHN
Swiss
Personalized
Health
Network

Borders exist, but are open



Current economic reality

SPARQL

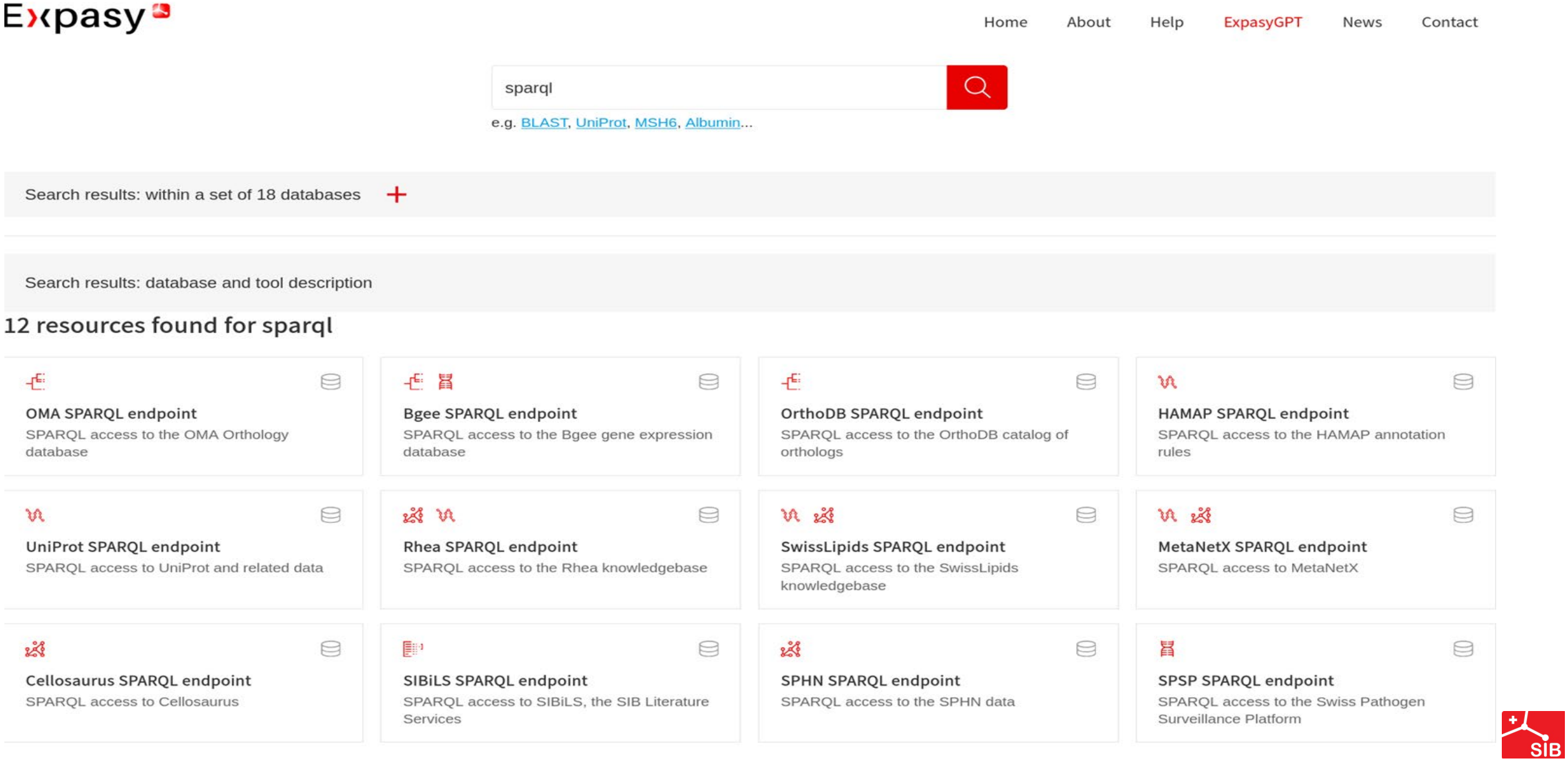
&

RDF

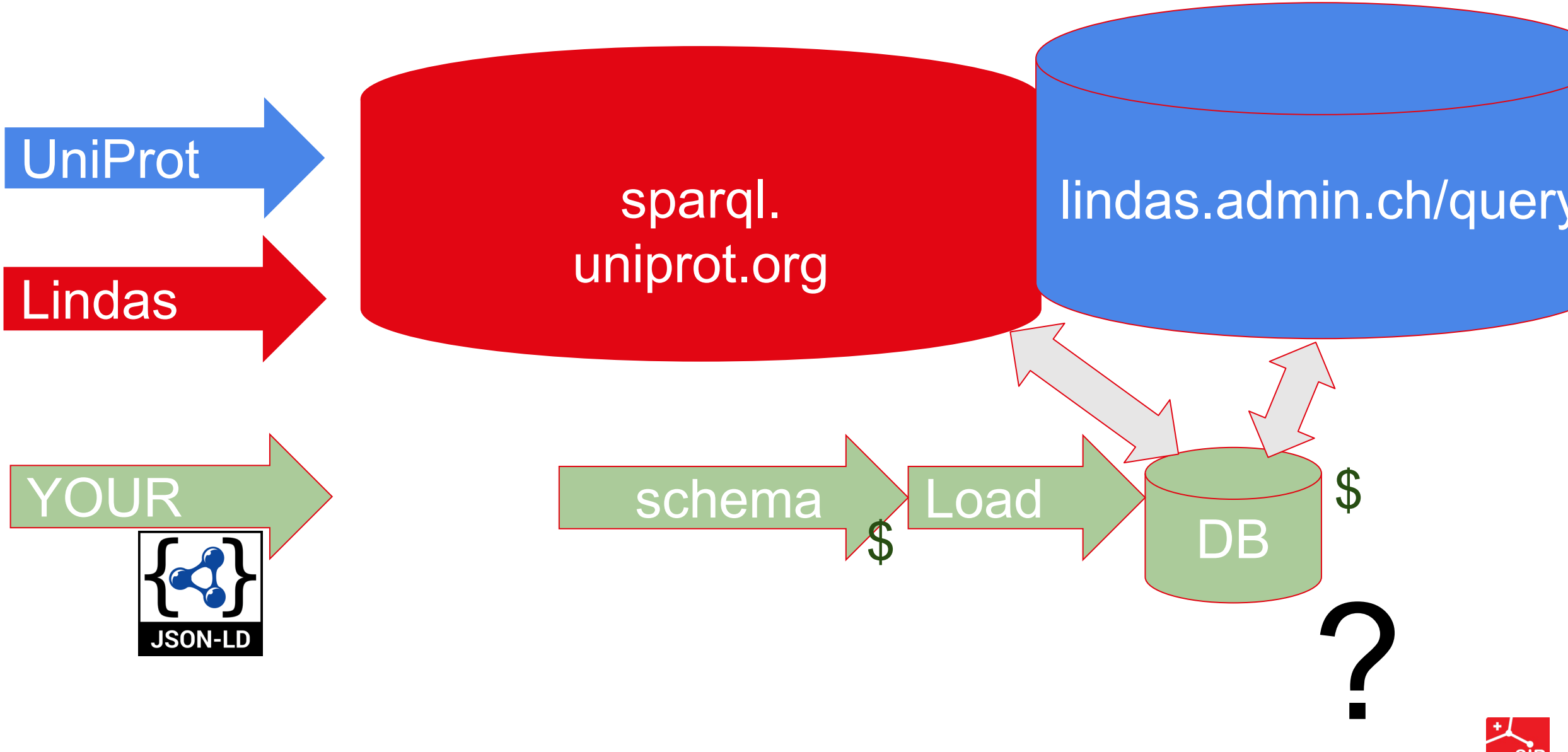
FAIR principle: interoperable **1**

data use a formal, accessible, shared, and broadly applicable language for knowledge representation

www.expasy.org/search/sparql



RDF + Federated SPARQL

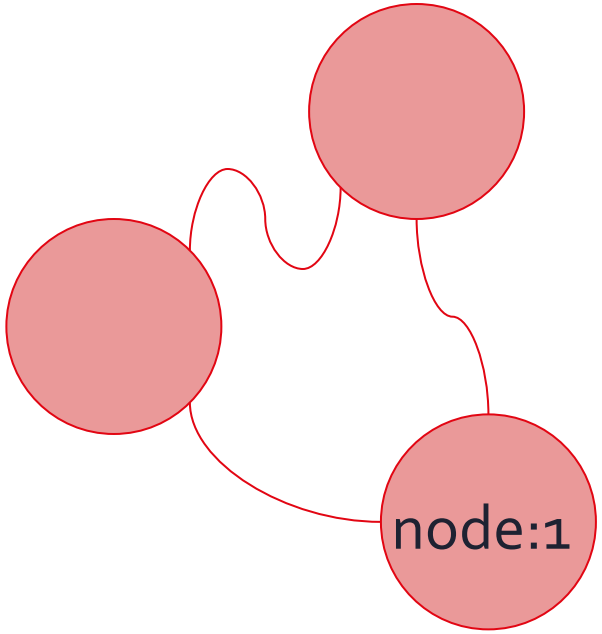


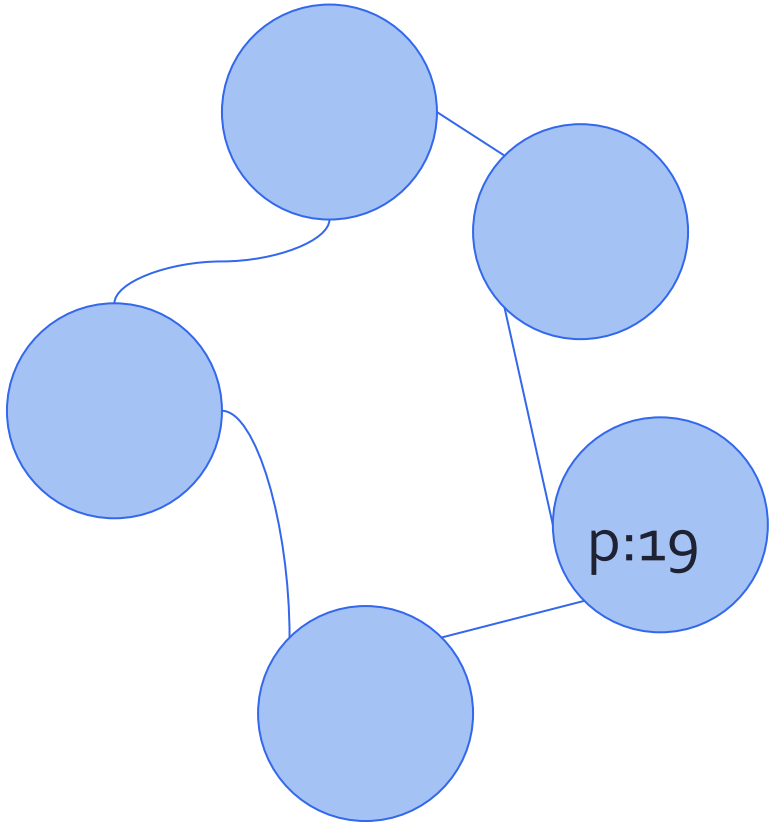
being

FAIR

allows this

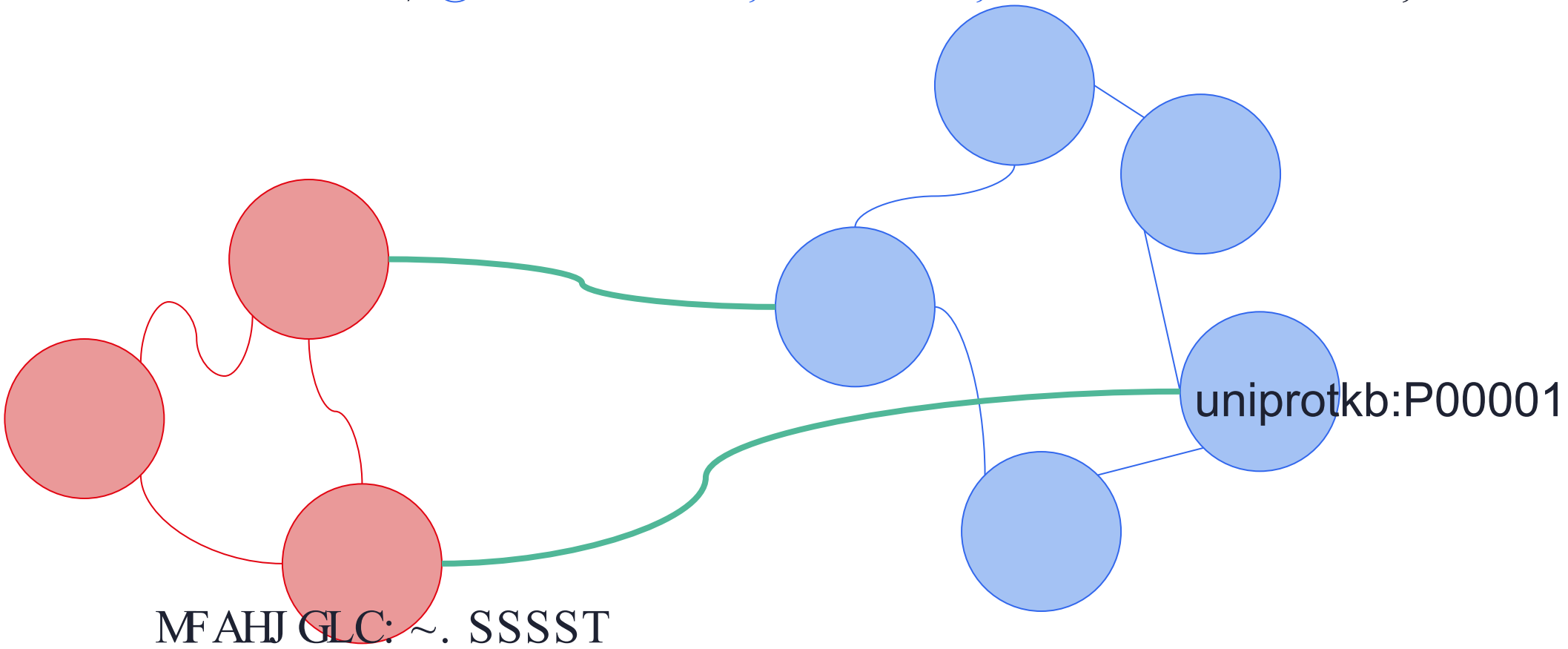
Just graphs are not the answer

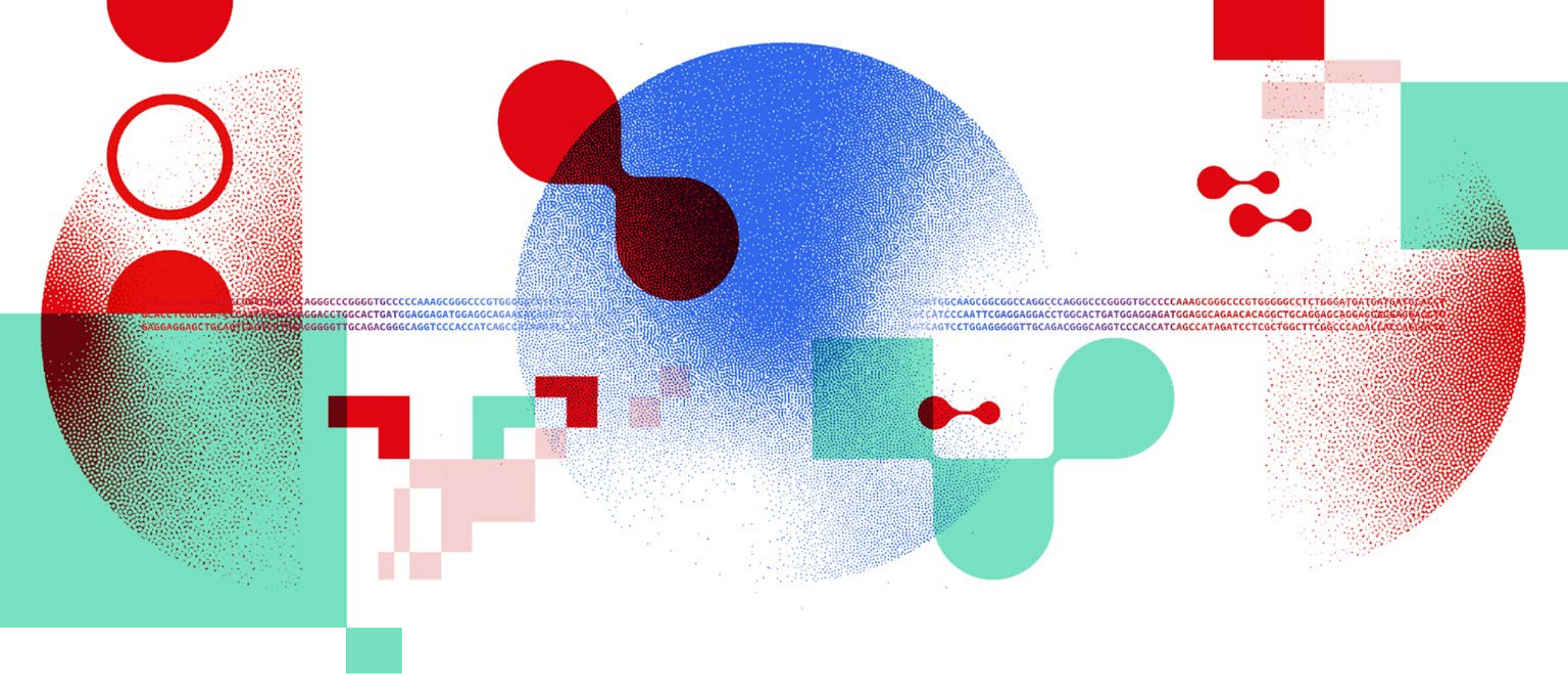




One **semantic** graph is

∑HJ =>AP MFAHJ GLC: ~¢@LLH~ßßHMJ D} MFAHJ GL} GJ ?ß MFAHJ GLß£ }





Thank you & all SIB members

DATA SCIENTISTS FOR LIFE

sib.swiss