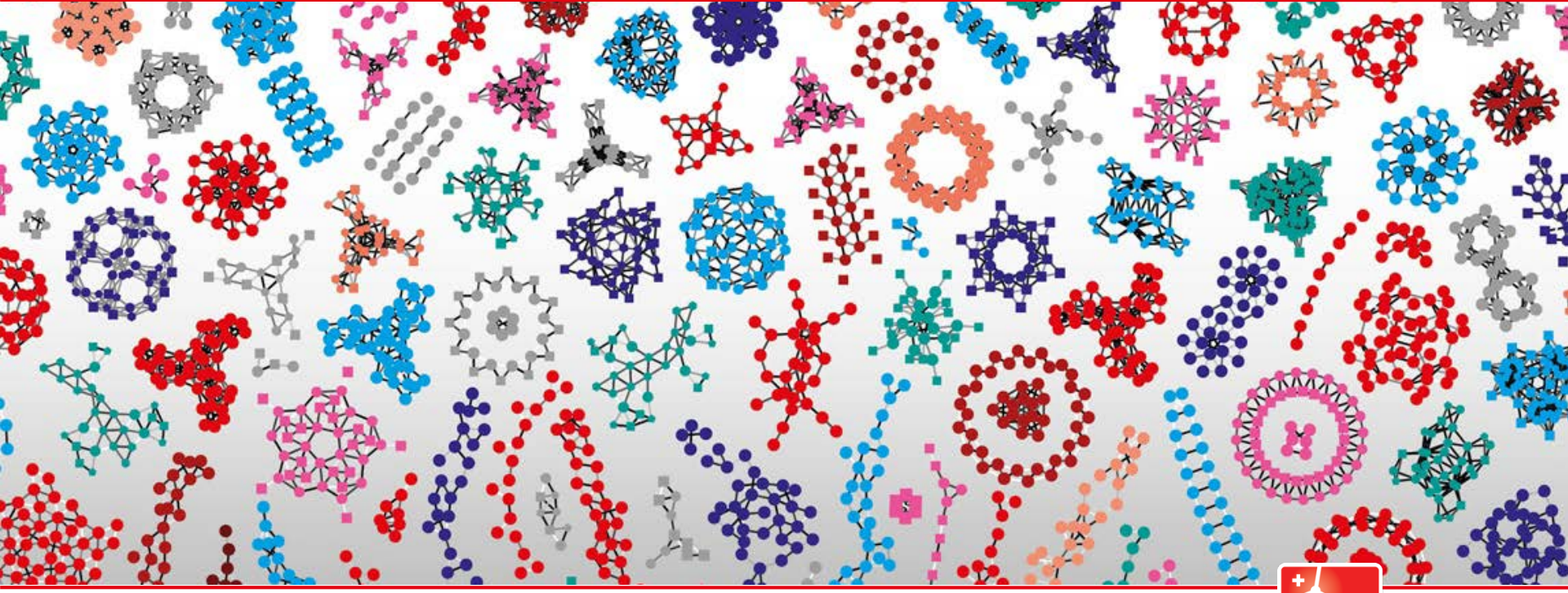
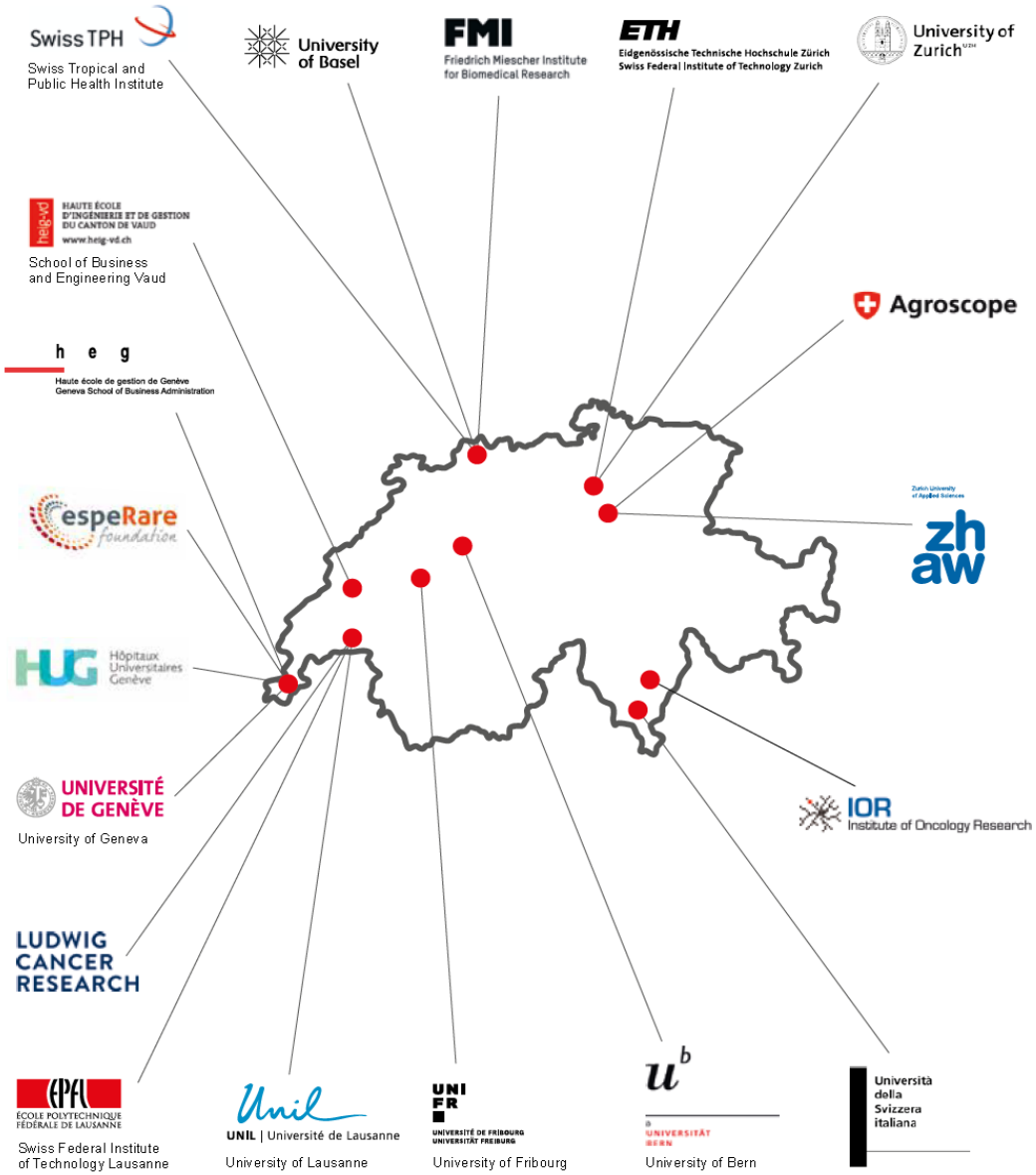


Swiss Institute of Bioinformatics

Sustaining open data resources in bioinformatics: the UniProt case study

Dr. Chiara Gabella, SIB Project Manager

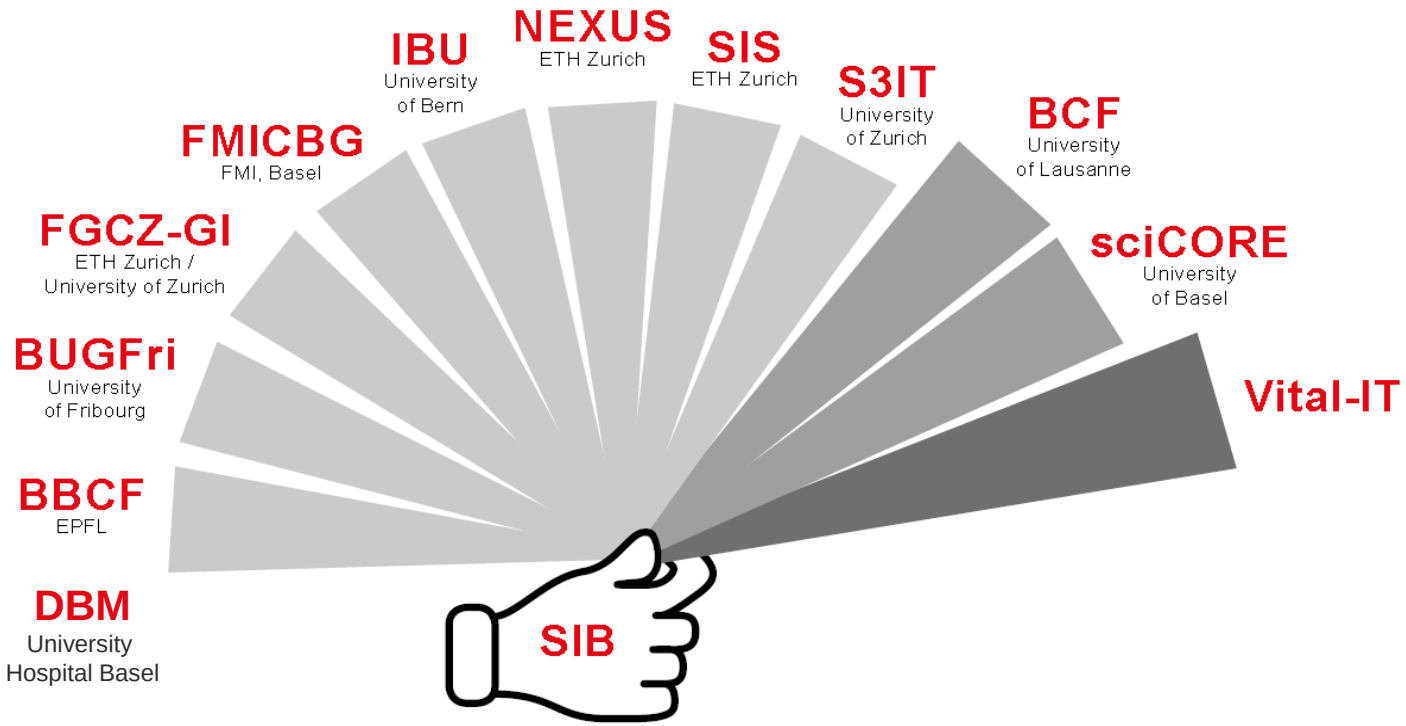
SIB Swiss Institute of Bioinformatics



19 partner institutions

65 groups
800 members

Bioinformatics competence centers



Resources for the international community

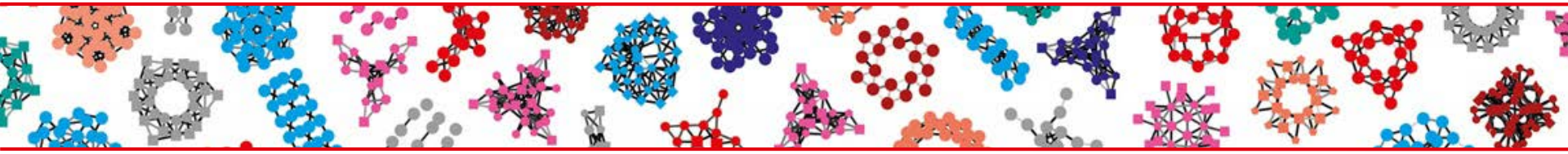
More than 150 bioinformatics resources

Resources for the international community

www.expasy.org

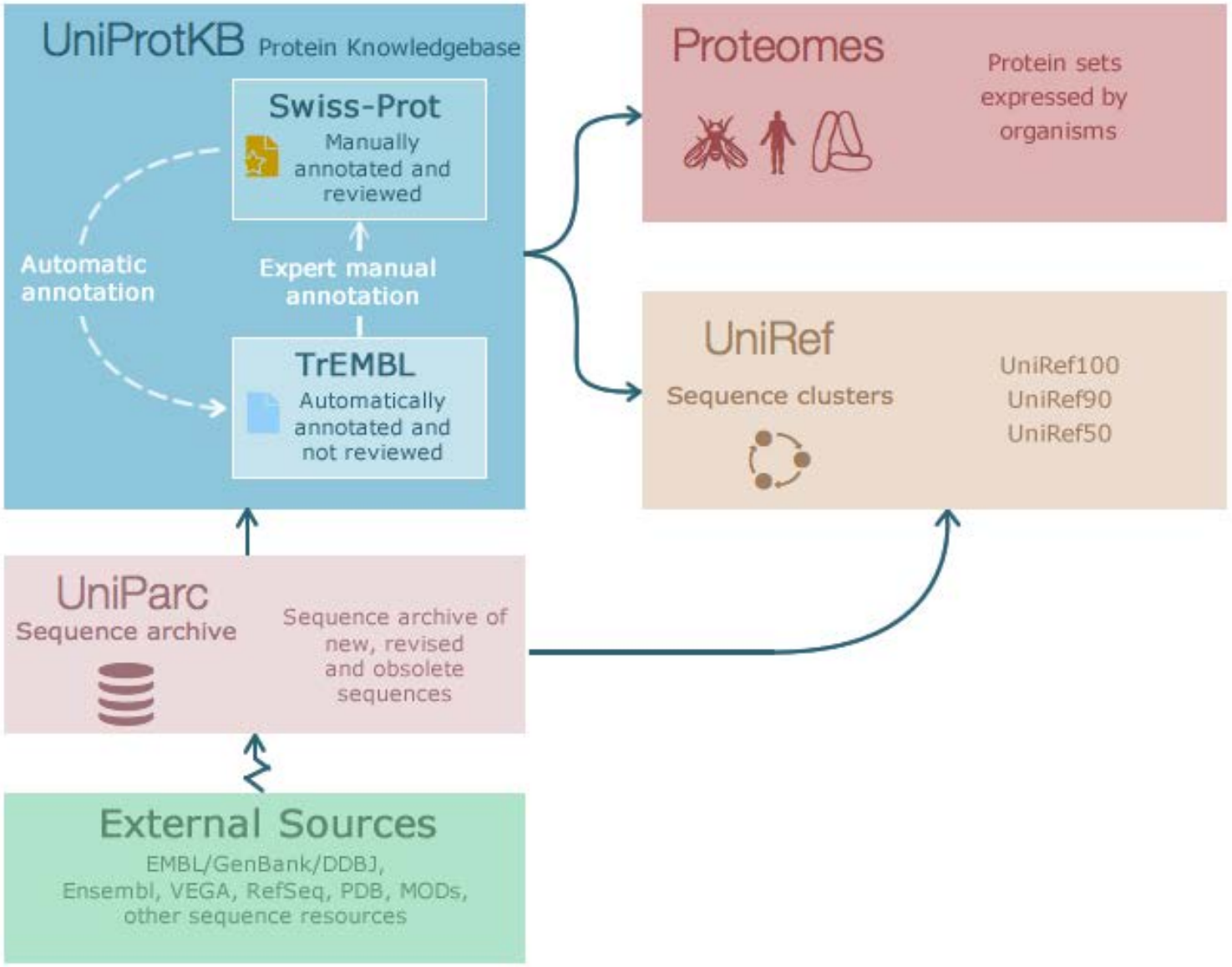
Knowledgebases

Expert manual curation



UniProt

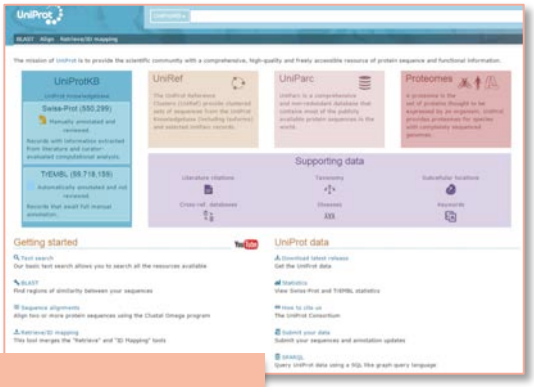
Resources for the international community



Resources for the international community

UniProtKB/Swiss-Prot

Encyclopedia of protein sequences and functional information

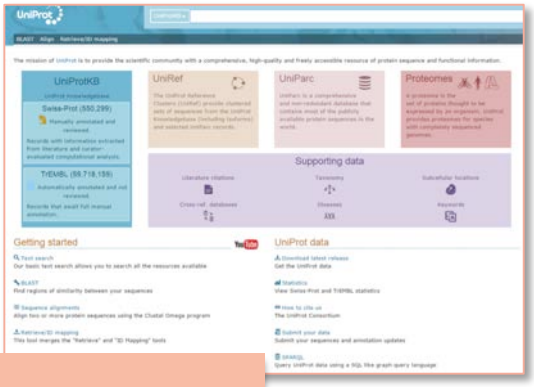


50 biocurators

Resources for the international community

UniProtKB/Swiss-Prot

Encyclopedia of protein sequences and functional information

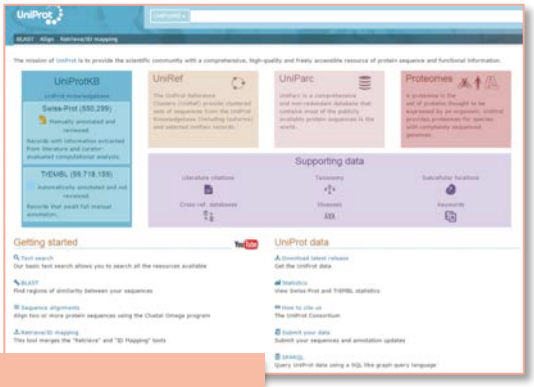


>550,000 manually curated protein sequences

Resources for the international community

UniProtKB/Swiss-Prot

Encyclopedia of protein sequences and functional information

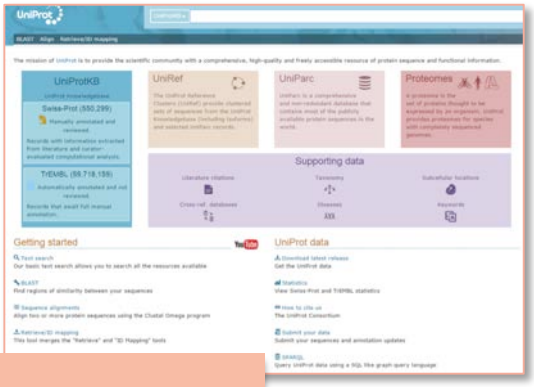


220,000 curated literature references

Resources for the international community

UniProtKB/Swiss-Prot

Encyclopedia of protein sequences and functional information



~500,000 users per month

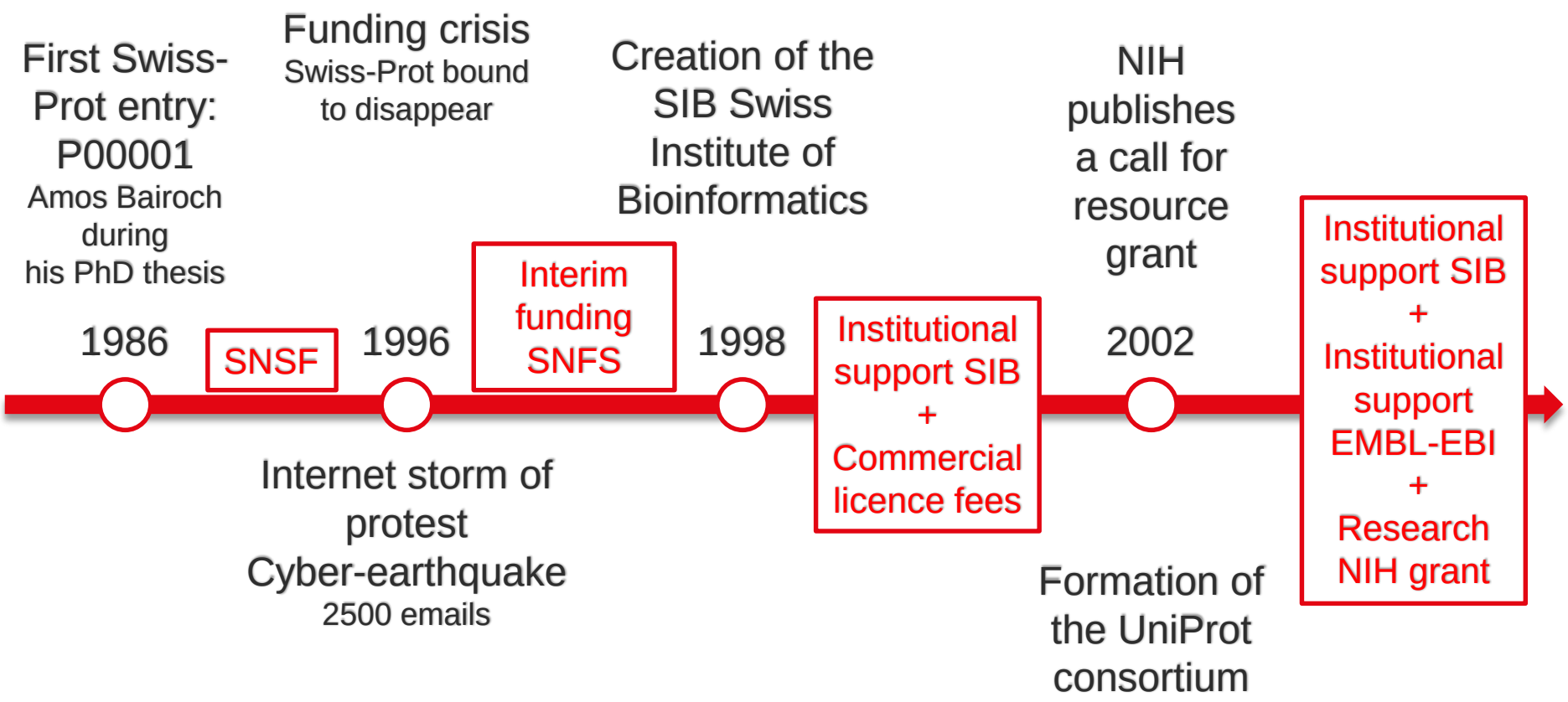
Access to UniProt is free

Cost of

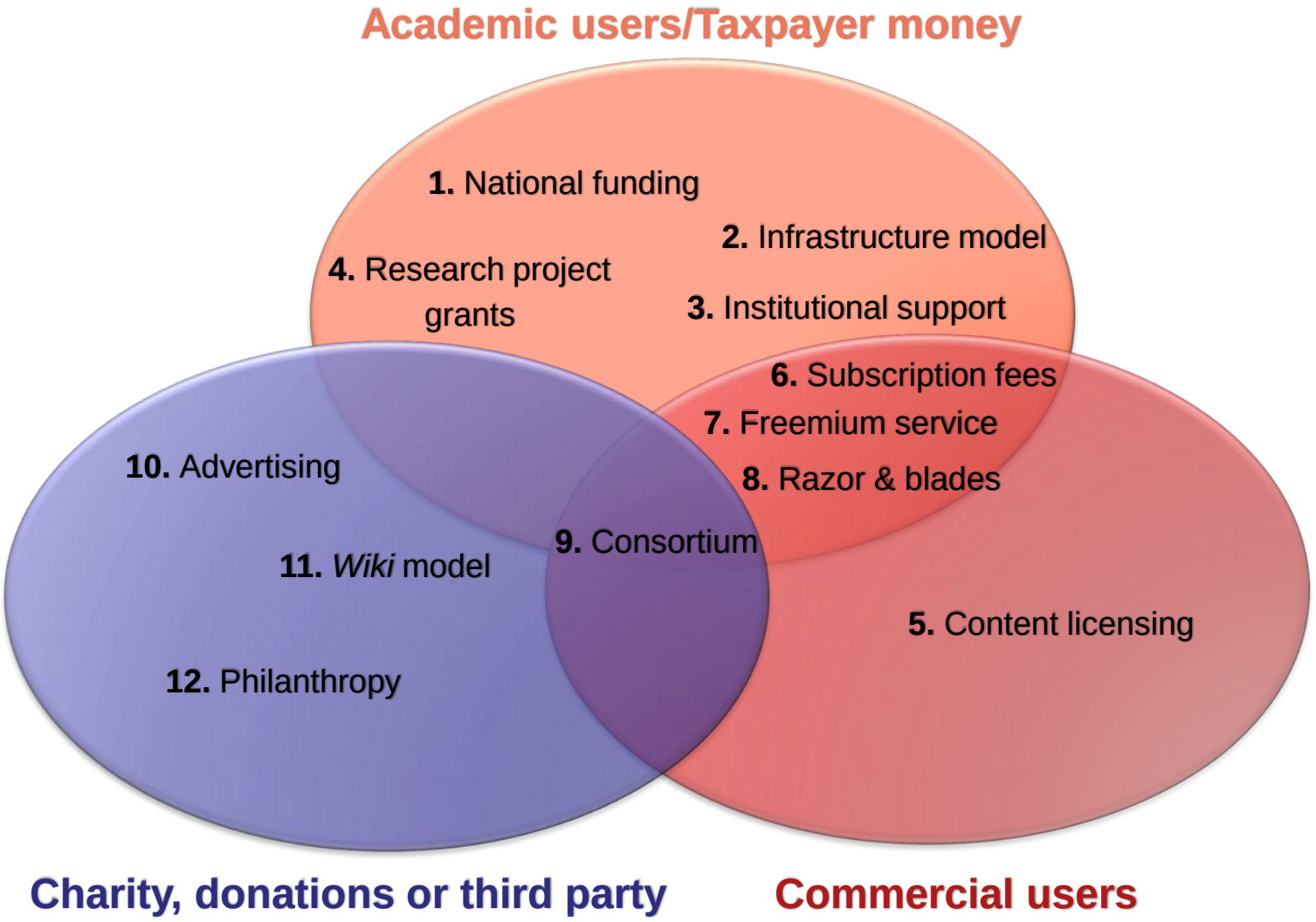
UniProt is
NOT
free

Cost of UniProt
≈ 0.1% of the
reviewed research

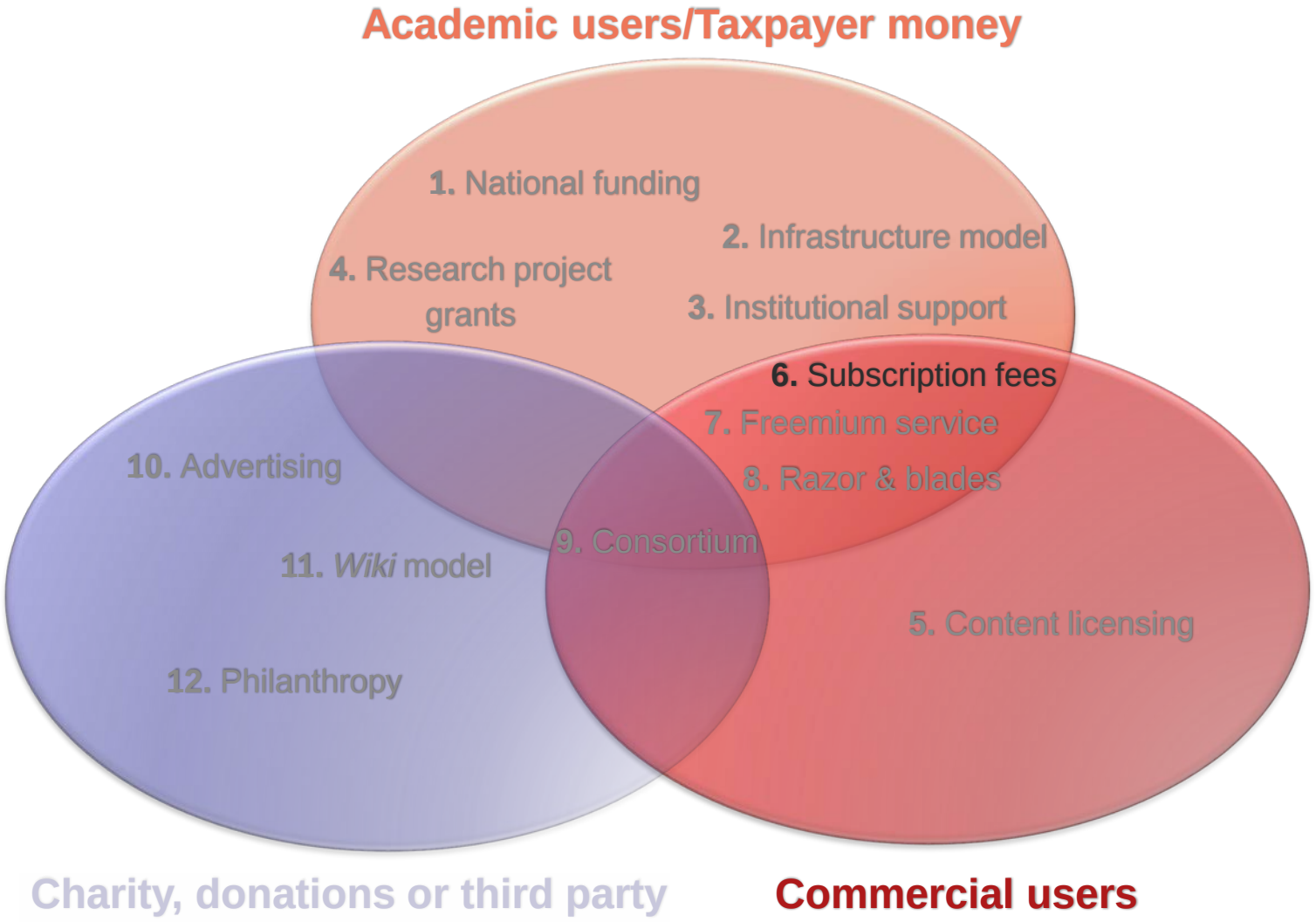
How is UniProt funded?



Possible funding models



Subscription fees



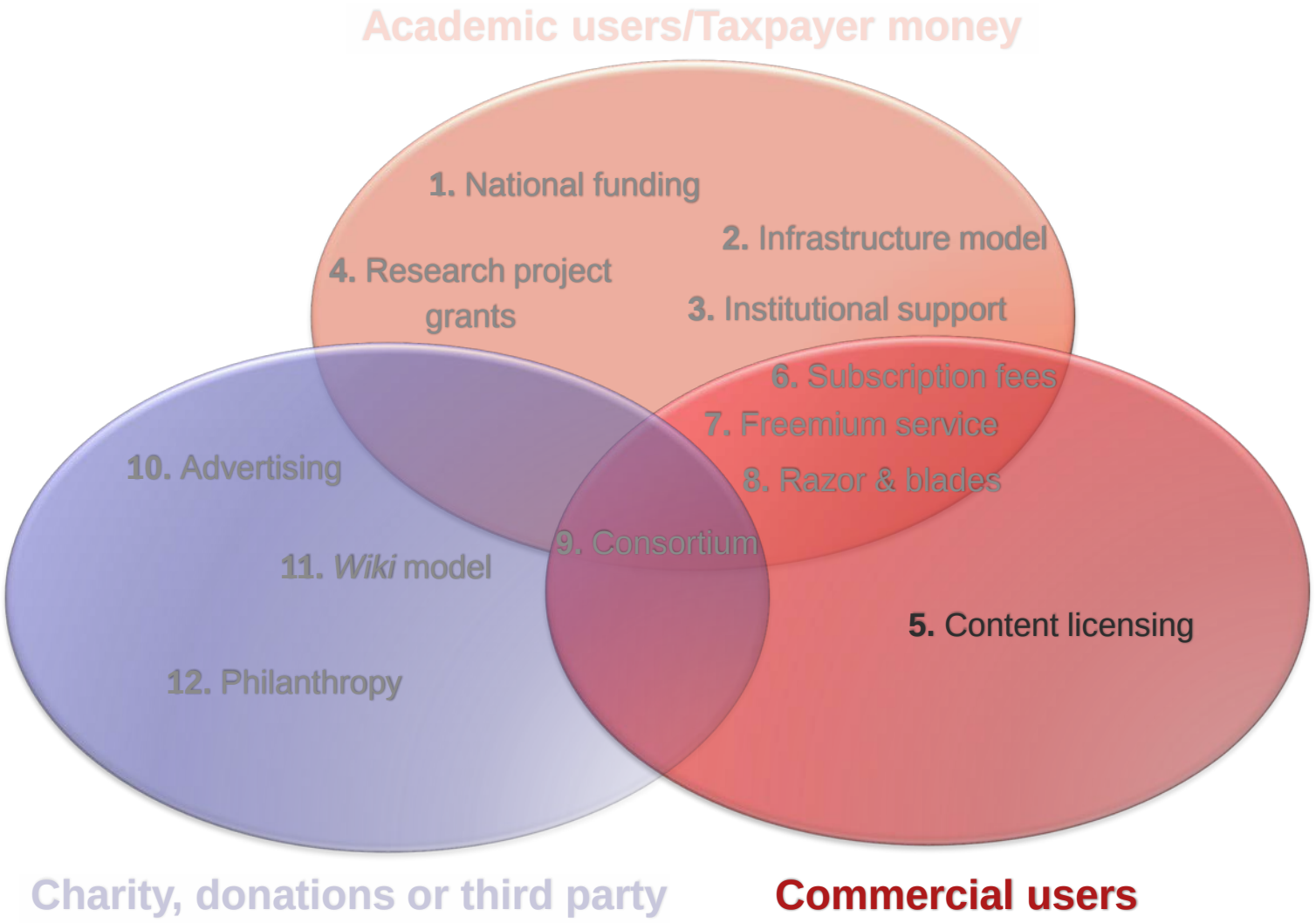
€ 3.5

/ month

€ 0.055

/ MB download

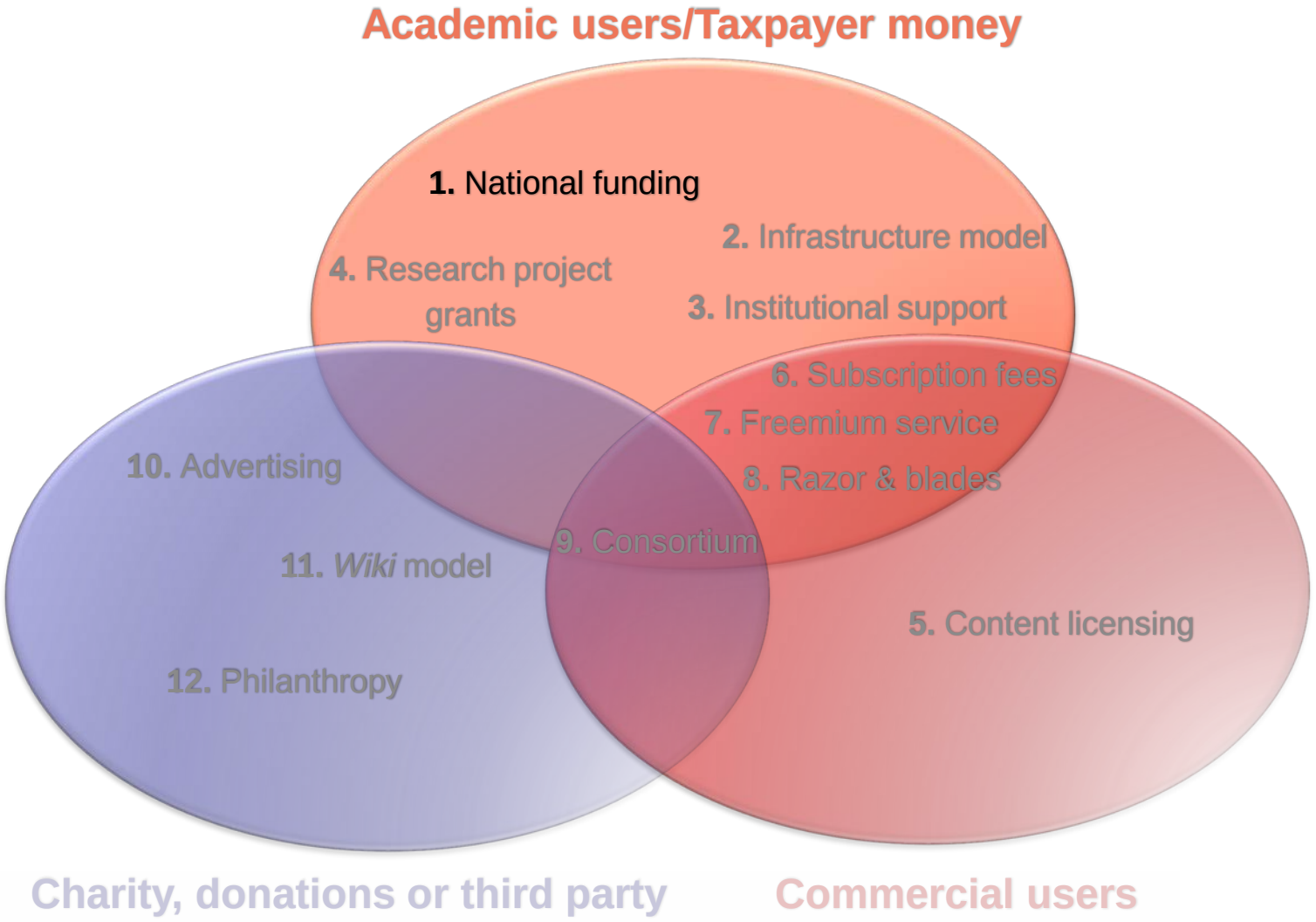
Content licensing

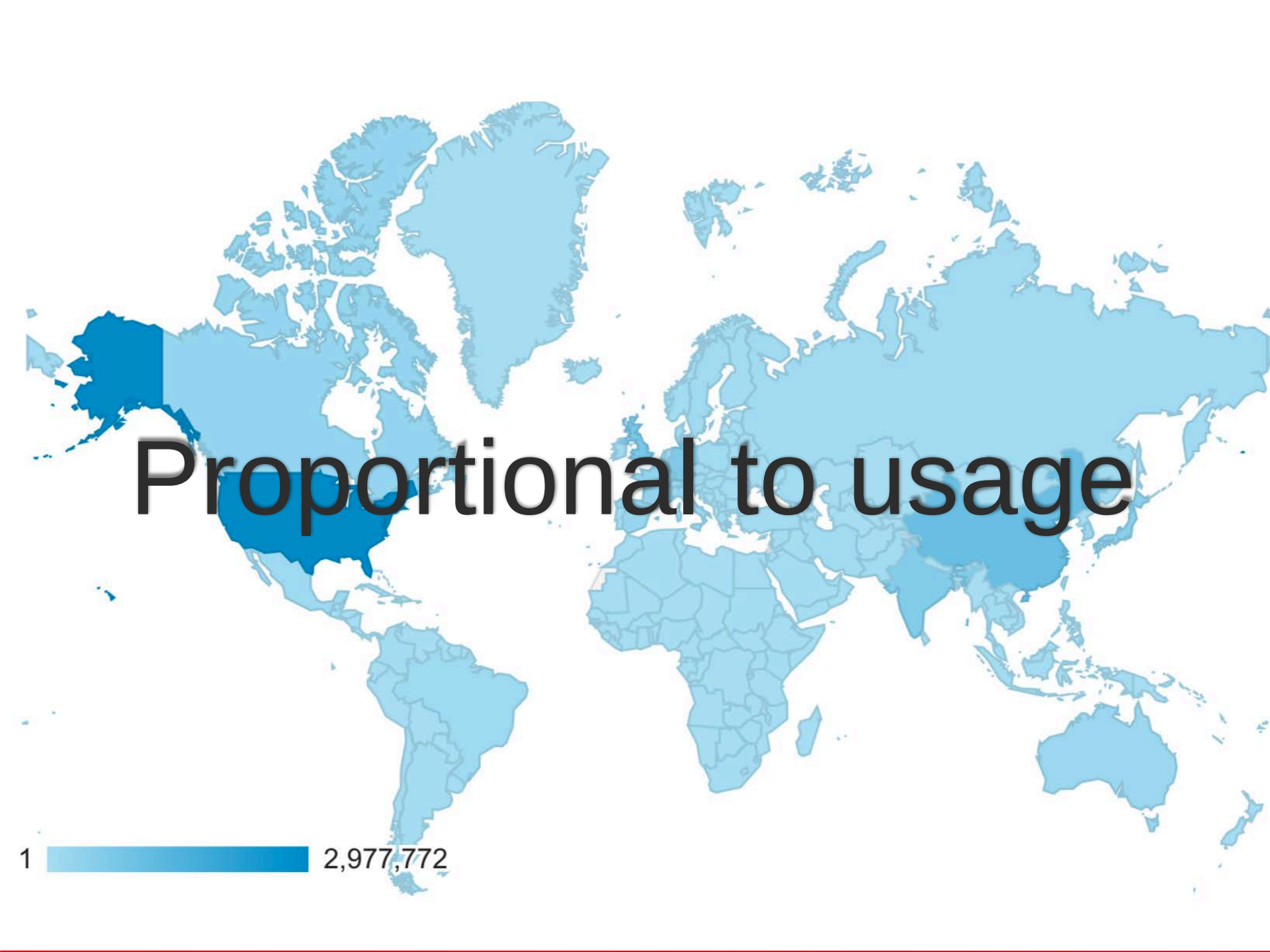


€ 2,500 - € 90,000

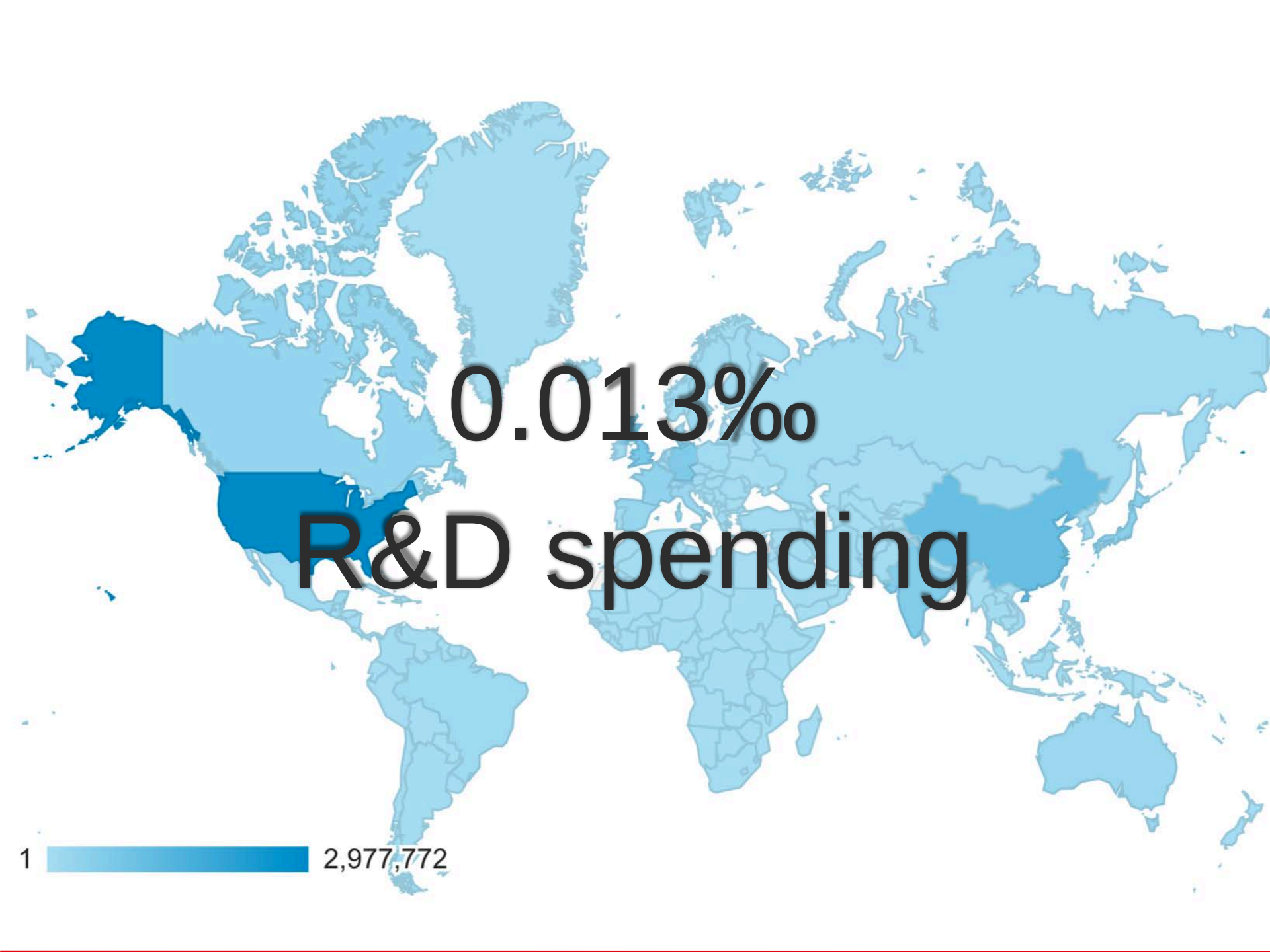
for Swiss-Prot subscription in 1998

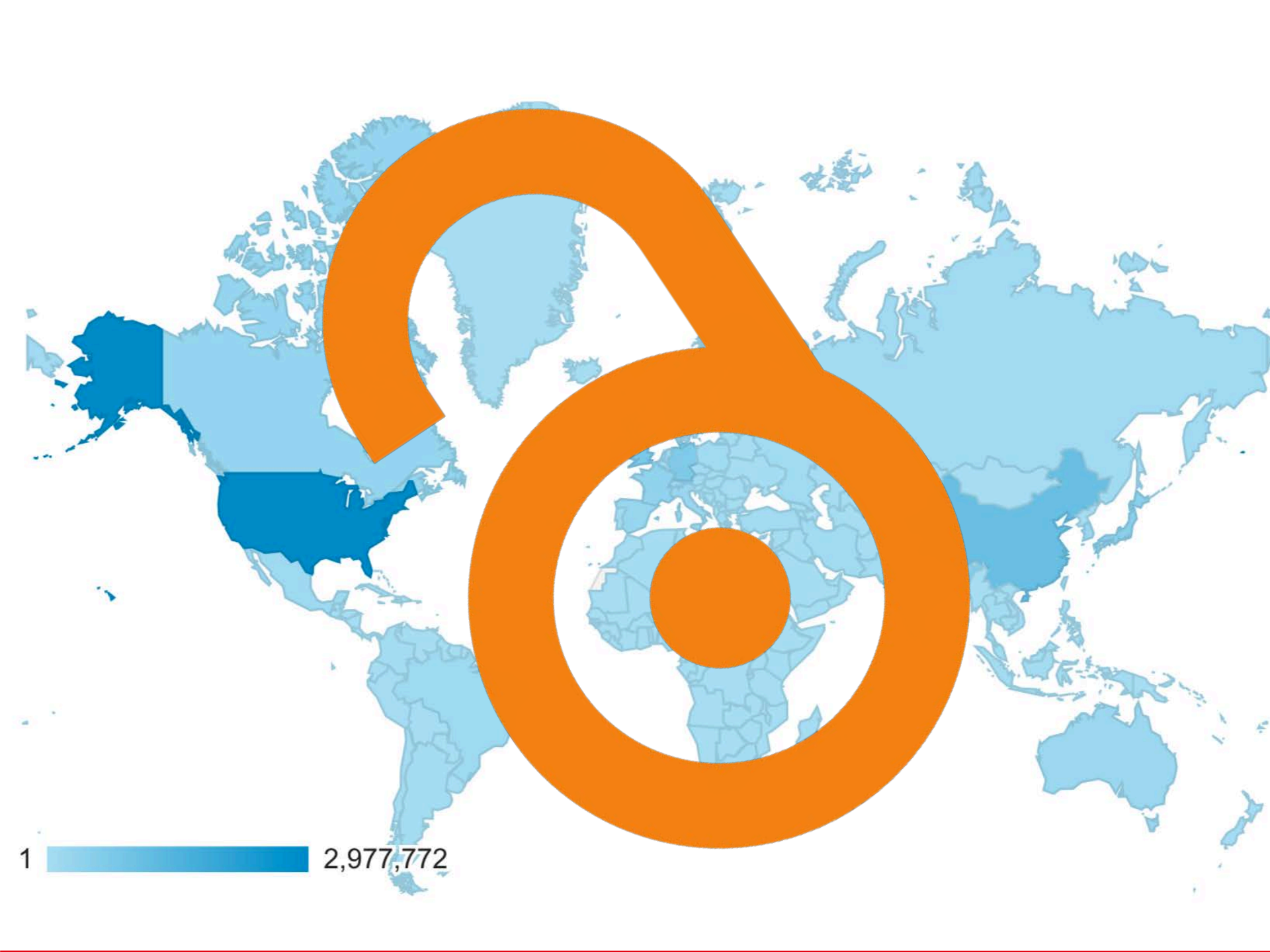
National model



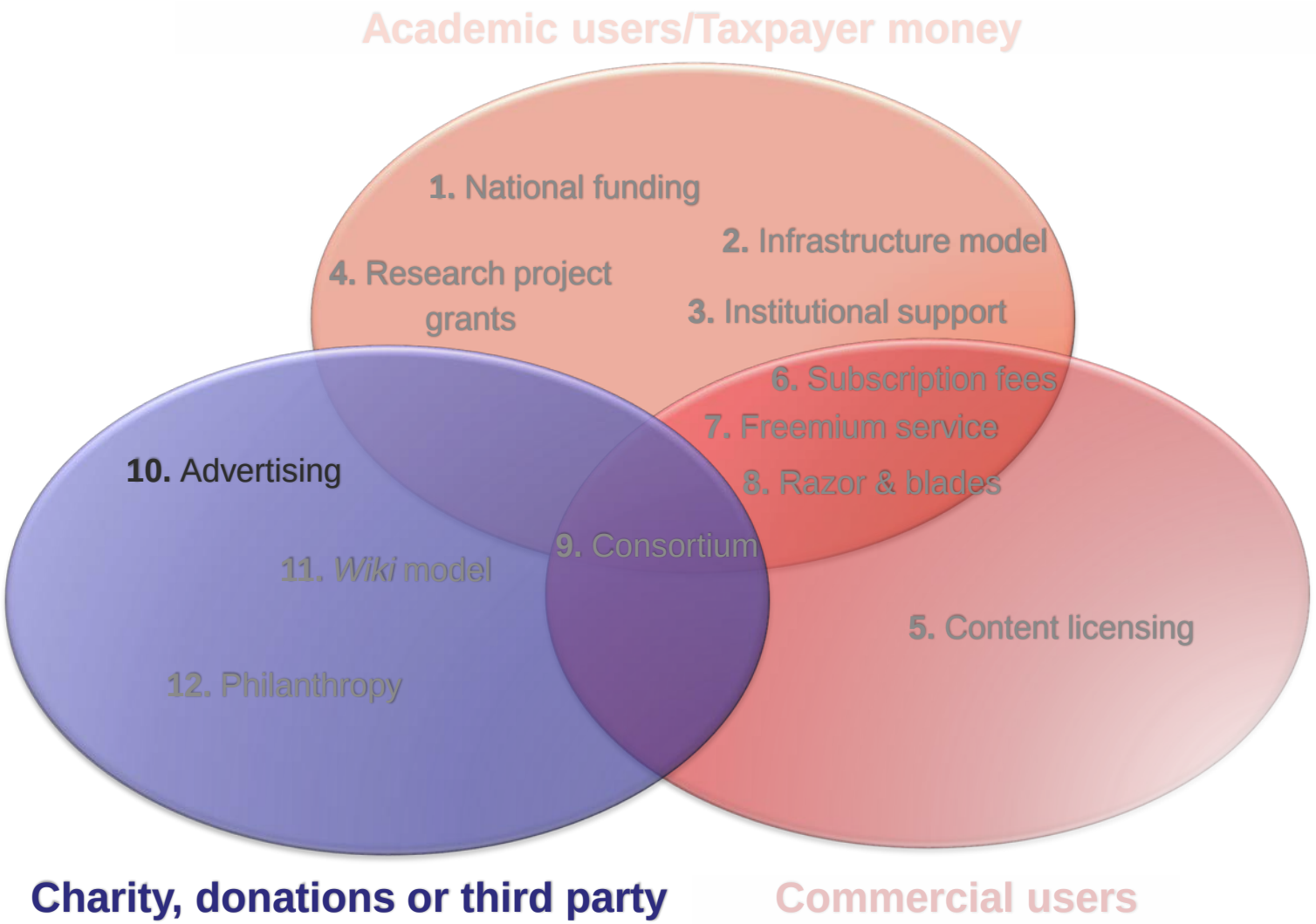

Proportional to usage
1
2,977,772

0.00025 – 0.00035‰
GDP
1
2,977,772

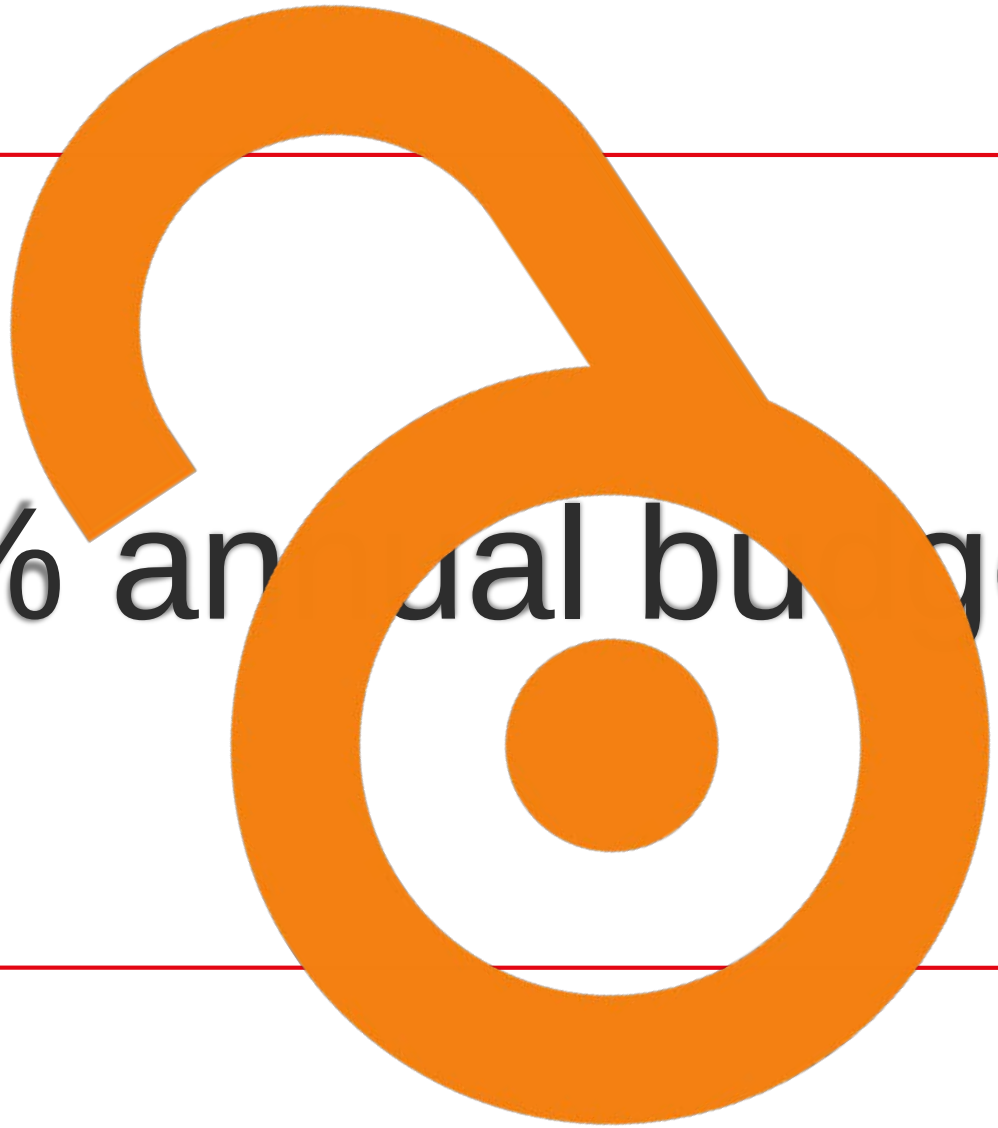

0.013‰
R&D spending
1
2,977,772


1
2,977,772

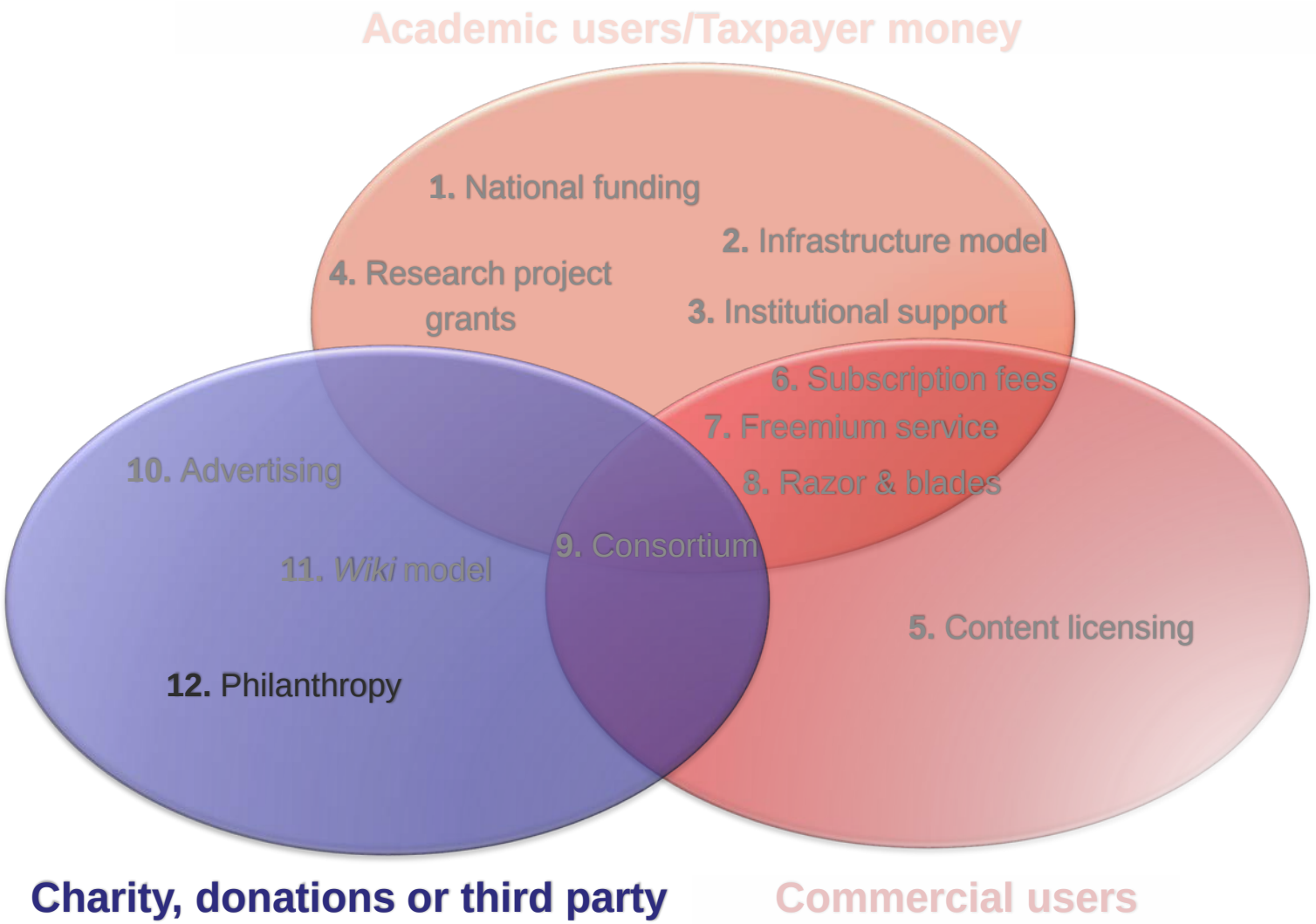
Advertising



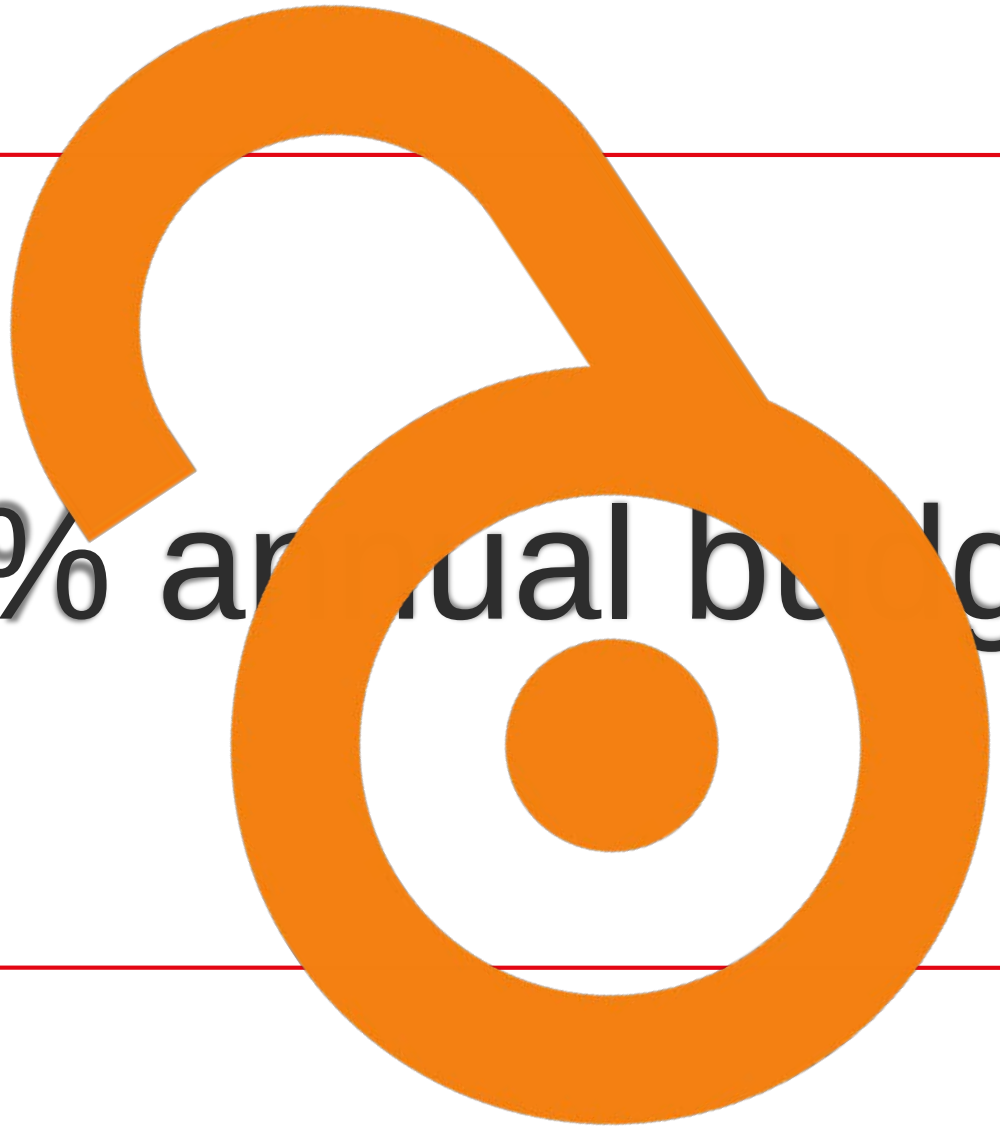
7% annual budget


7% an dal bu get

Philanthropy

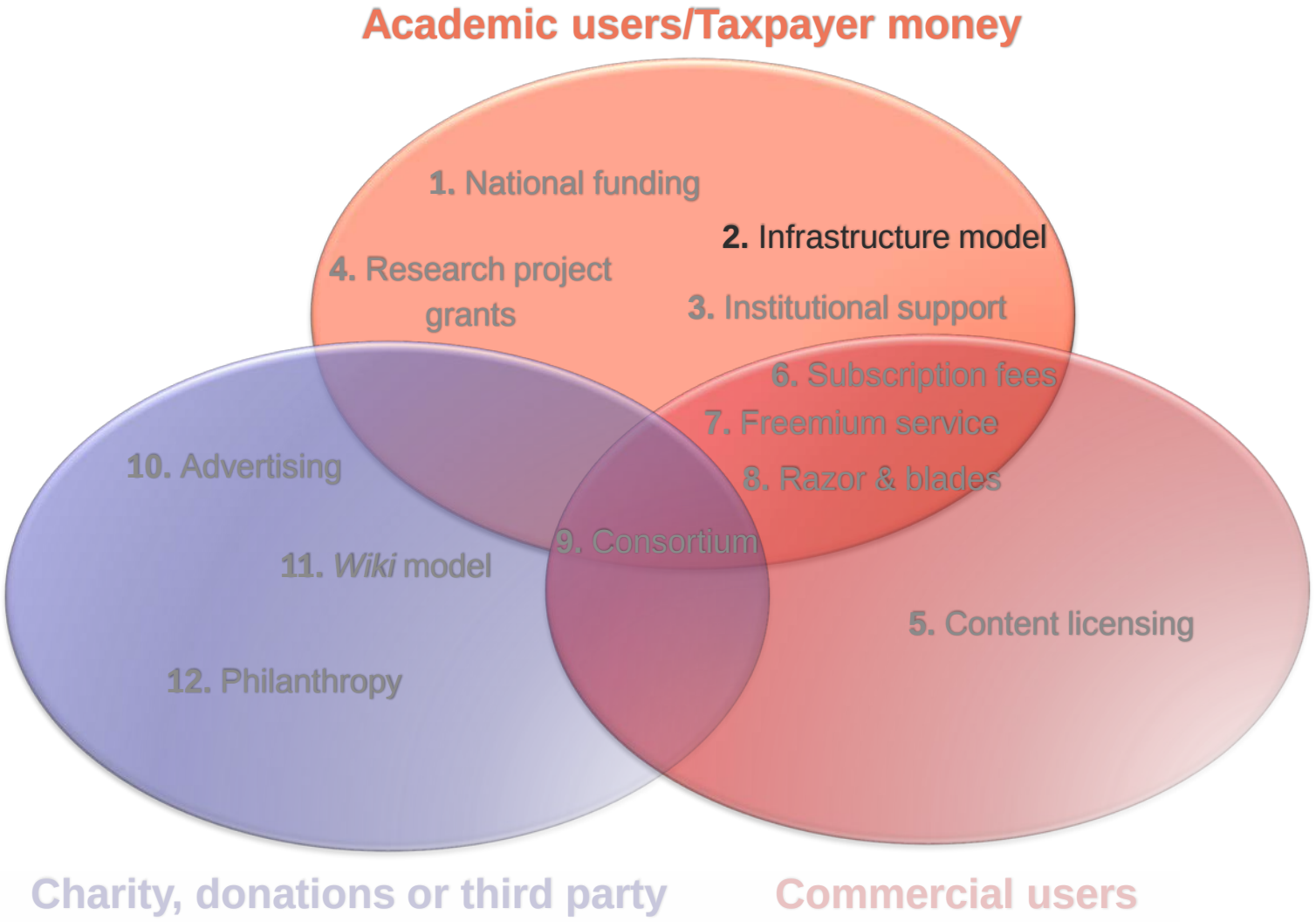


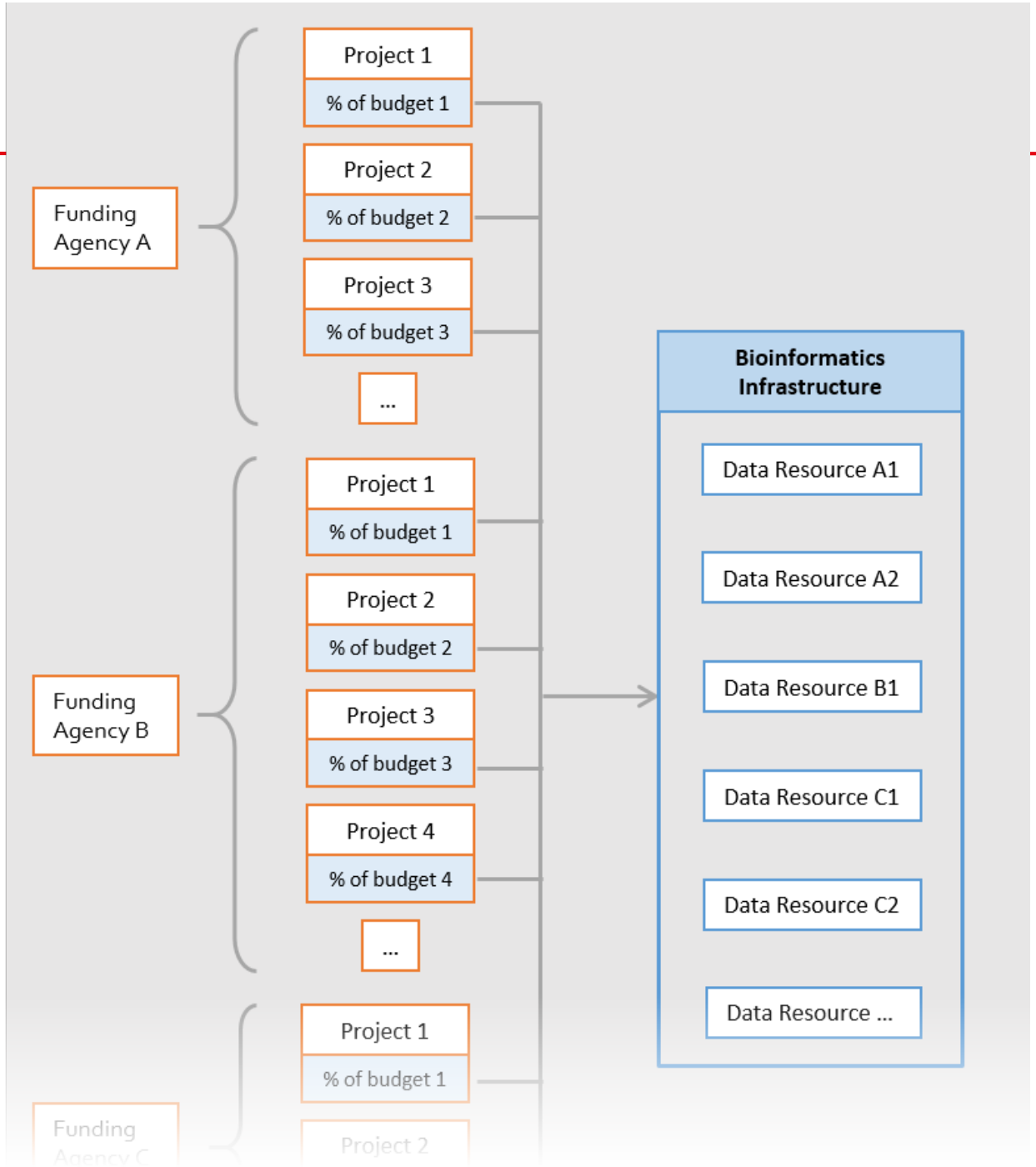
<5% annual budget



<5% annual budget

Infrastructure model



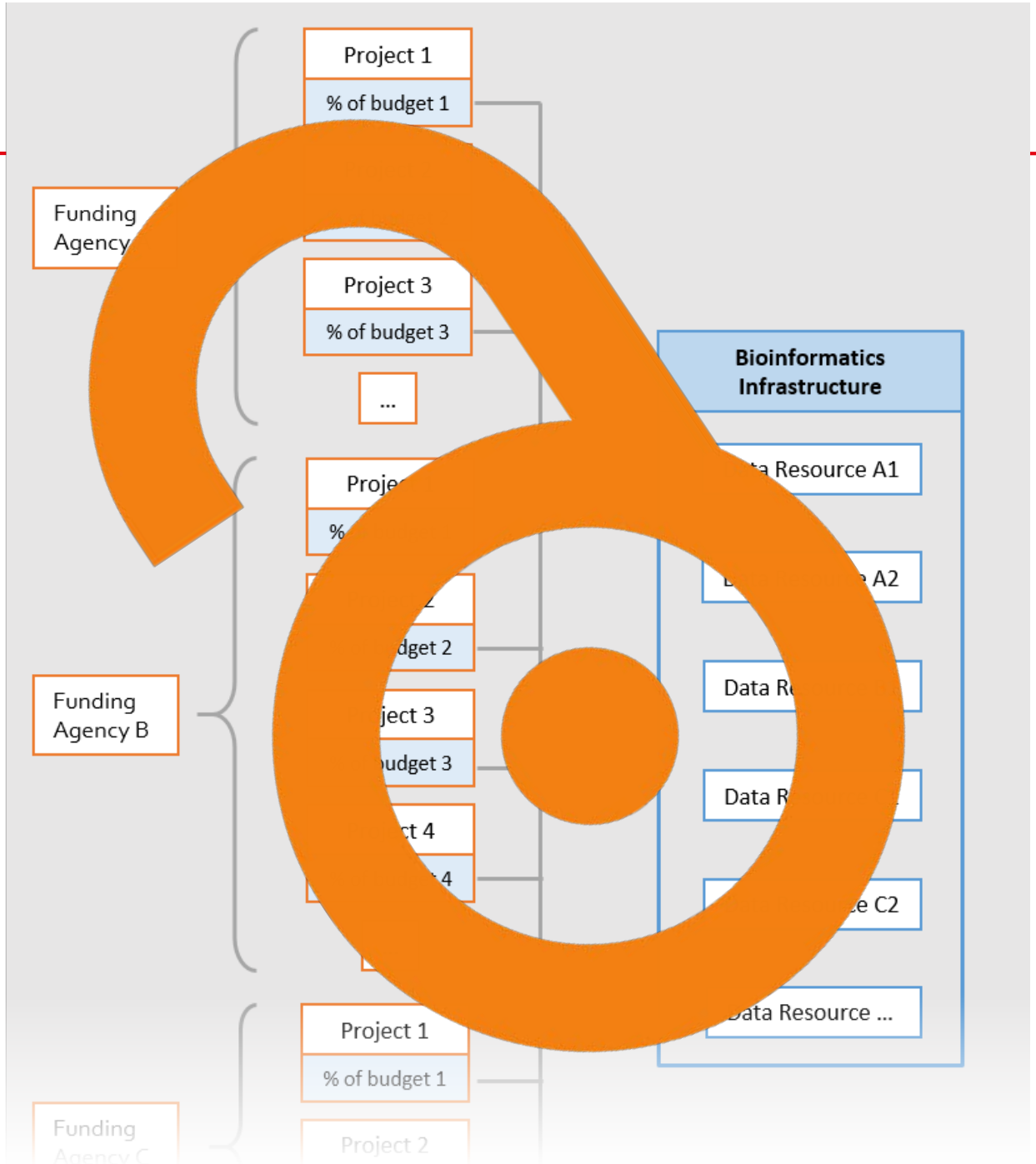

Funding Agency A
Project 1
% of budget 1
Project 2
% of budget 2
Project 3
% of budget 3
...
Funding Agency B
Project 1
% of budget 1
Project 2
% of budget 2
Project 3
% of budget 3
Project 4
% of budget 4
...
Funding Agency C
Project 1
% of budget 1
Project 2
Bioinformatics Infrastructure
Data Resource A1
Data Resource A2
Data Resource B1
Data Resource C1
Data Resource C2
Data Resource ...

5 funding agencies

0.25 %

total spending for life science research grants

could sustain UniProt


Project 1
% of budget 1
Funding Agency
Project 3
% of budget 3
...
Bioinformatics Infrastructure
Resource A1
A2
Funding Agency B
Data R
Data R
C2
Data Resource ...
Project 1
% of budget 1
Project 2

Scalable with the research

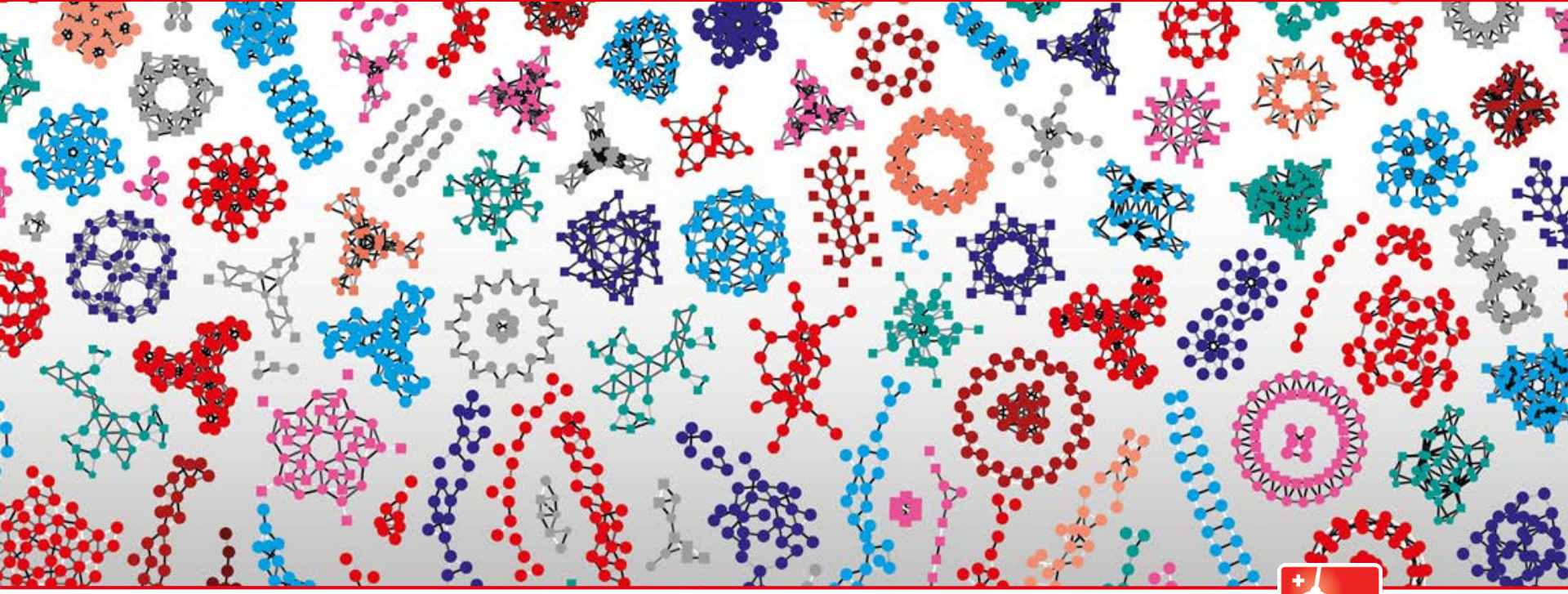
International

Open data

is not free

data

Long term sustainable model is needed



SIB
Swiss Institute of Bioinformatics

Thank you

elixir SWITZERLAND

www.sib.swiss