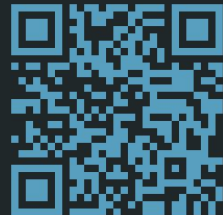




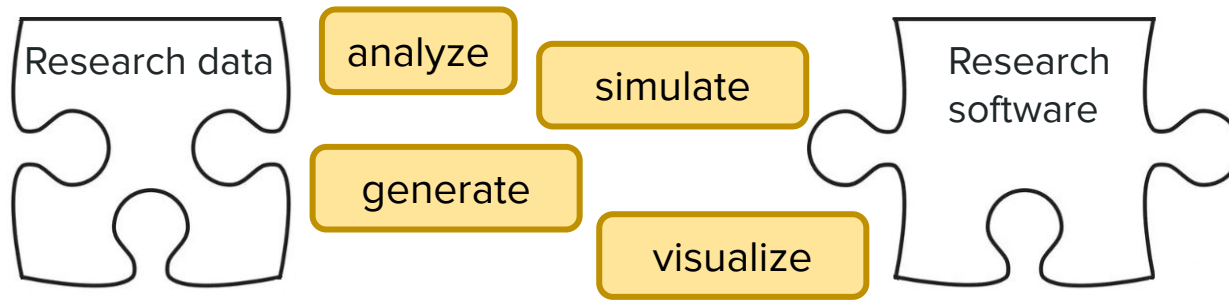
UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

What institutions can do to provide support

Inga Ulusoy
Scientific Software Center, Heidelberg University



The digital transformation of scientific and scholarly research: Institutional support



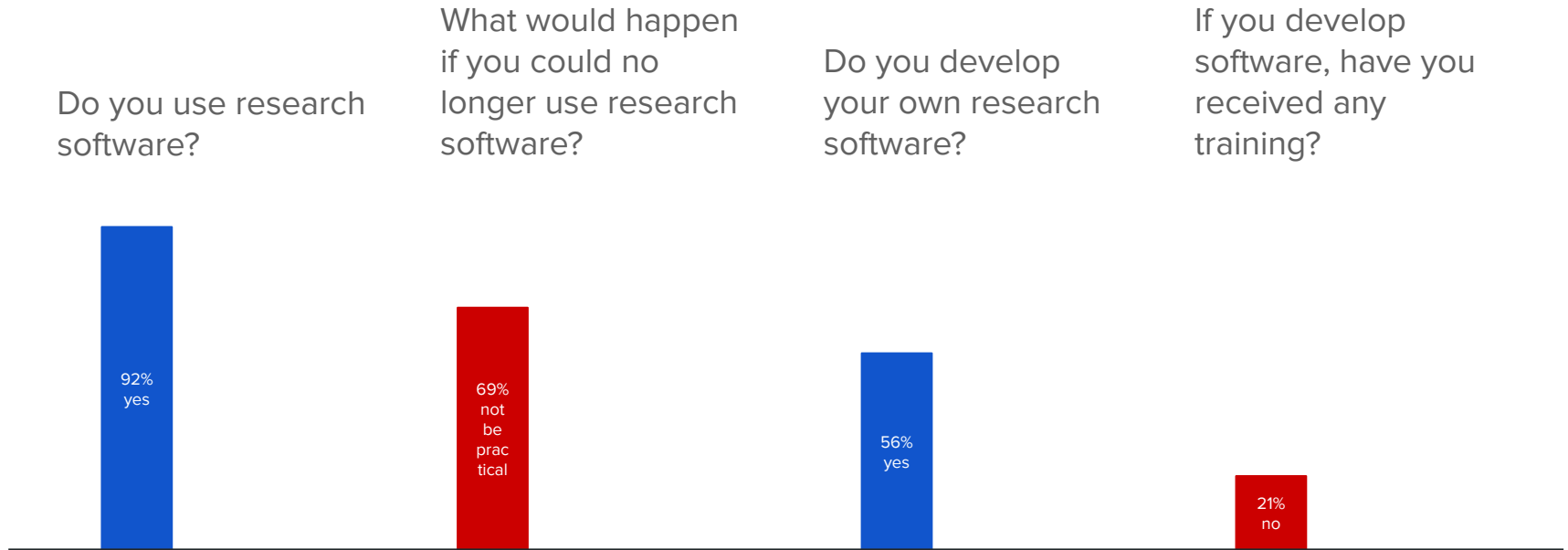
Research Software

- Computer program developed and used in the context of scientific research
- Serves specific purpose
- Requirements determined in scientific process

Research Software Engineer (RSE)

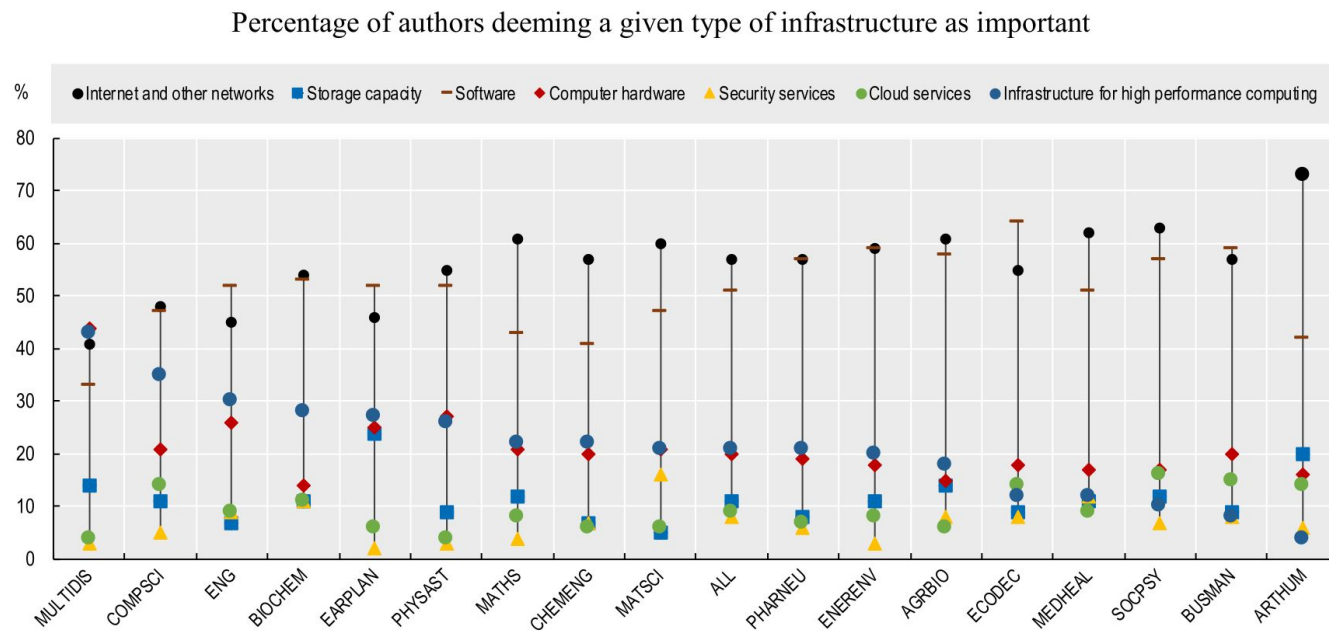
- Researcher that is familiar with both the scientific process and software engineering

Use and development of research software



Research infrastructure

Figure 5.5. Most important infrastructure for scientific authors' research work, by field



Note: Weighted estimates based on sampling weights adjusted for nonresponse. Respondents can select a maximum of two options.

Source: OECD International Survey of Scientific Authors (ISSA), 2018. <http://oe.cd/issa>.

Objectives

- Provide institutional support: Research software services as infrastructure
- Support researchers through training in software engineering

Scientific Software Center (SSC): Mission statement

Improve scientific software development practices to promote reproducible science and research software sustainability

- provide a research software infrastructure to support research
- provide development support for researchers that develop their own research software
- provide training for researchers that develop their own research software
- promote research software and their re-use
- support and promote the recognition of research software as a scientific output
- support the recognition of research software engineering as viable academic career option
- promote and practice open-source access and sharing of research software
- foster synergies between scientific and scholarly disciplines

Support offered by the SSC: Development & Sustainability

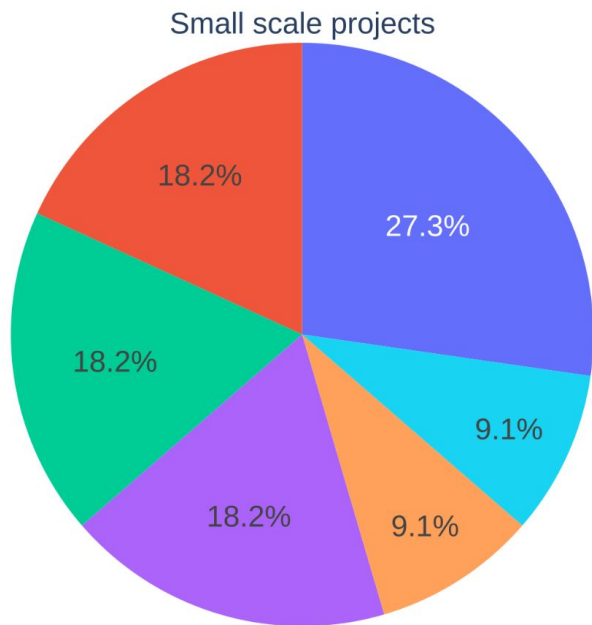
Target
Group leaders

Aim

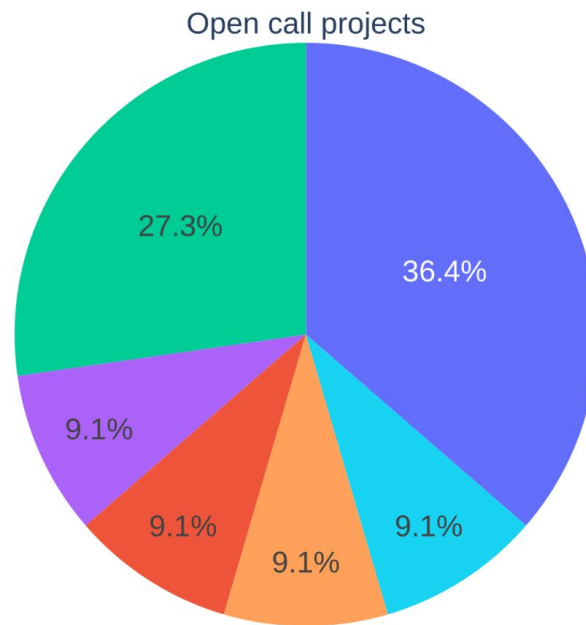
- Support researchers with high-quality, sustainable research software that generates reproducible data
- Provide access to well-trained RSEs
- Provide a career option and a community to RSEs
- Adapt the latest software engineering standards in research software development

<i>type of support</i>	small-scale project	open call project	third-party funded project
<i>time scale</i>	1 week	1 – 12 months	> 1 month
<i>cost</i>	free of charge	free of charge	determined through consultation
<i>availability</i>	throughout the year	proposal submission deadline mid-June	throughout the year depending on availability of resources

Distribution of small-scale and open call projects 2021-2023



■ Life Sciences
■ Natural Sciences



■ Computer Science/Mathematics
■ Digital Humanities
■ Social Sciences
■ Earth Sciences

Example: Small-scale project

Hammingdist

Topological data analysis of the genomic structure of variants of the SARS-CoV2-virus (Physical Mathematics)

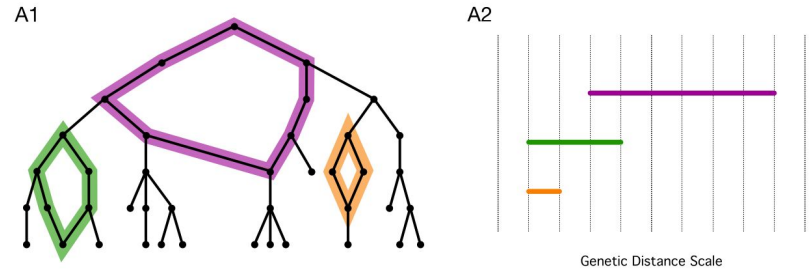
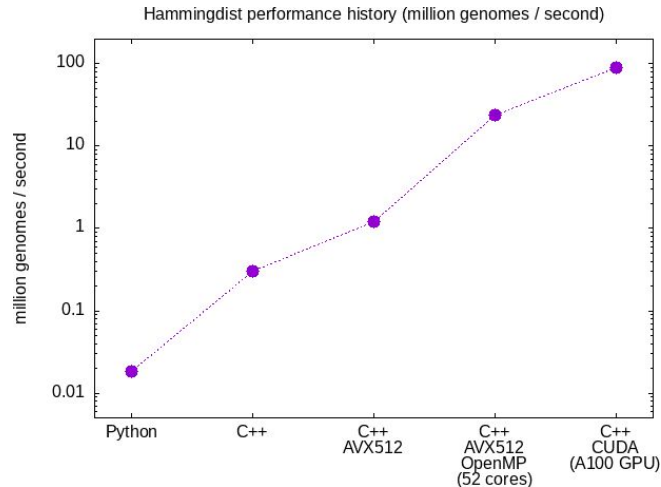


Figure 2. Topological data analysis quantifies convergent evolution.

Compute genetic distance scale of reticulate events in viral evolution.

<https://arxiv.org/pdf/2106.07292.pdf>



Improved implementation currently handles 10 GB of genome data ($\approx 350k$ genomes each of length $\approx 30k$).

<https://github.com/ssciwr/hammingdist>

Example: Open call project ammico

AI Media and Misinformation Content Analysis Tool (Social/Political Sciences)



<https://github.com/ssciwr/AMMICO>
<https://osf.io/preprints/socarxiv/v8txj>
<https://doi.org/10.57967/hf/0603>

TextDetector x

Analyse text

Select models for
text_summary,
text_sentiment,
text_NER or leave
blank for default:

Select model
revision number for
text_summary,
text_sentiment,
text_NER or leave
blank for default:

Run Detector

filename	../data/Image_some_text/1092375_spa.png
text	29 de septiembre CONFÍAN EN LA REUNIÓN DE HOY 0:10/0:14
text_language	es
text_english	September 29th THEY TRUST IN TODAY'S MEETING 0:10/0:14
text_clean	September 29th THEY TRUST IN TODAY'S MEETING
text_summary	September 29th THEY TRUST IN TODAY'S MEETING 0:10/
sentiment	POSITIVE
sentiment_score	0.99
entity	
entity_type	

Support offered by the SSC: Teaching & Consultation

Target

PhD + Master students + Postdocs for training
All of these + group leaders for consultations

Aim

- Provide access to training in software engineering best practices
- Spread the use of best practices in the community
- Provide an easy, low-barrier entrance point for researchers, directly provide feedback on software quality or software-related questions

Block
courses

On-demand
courses

Guidelines

Mentoring
program

Consultations

Template
repositories

Compact
courses

Seminars

Example: Compact/block courses

The Unix Shell

Version Control with git

Open Source Licensing

Automated Testing with GitHub Actions

A short tour of sustainable software development

Containers in Science: Using Docker and Singularity

Advanced Topics in Version Control with git

Effective Software Testing

Python Best Practices

Introduction to Python Testing

Data Exploration with Python and Jupyter

Python Packaging

Scientific Software Development

Lunch-time Python

Performance Benchmarking C++ Applications
High Performance C++

Support offered by the SSC: Outreach & Communication

Target

National initiatives

Funding agencies

Local institutions

Aim

- Raise awareness for software engineering best practices
- Improve recognition of research software as scientific output
- Support and shape RSE career paths
- Support and enhance recognition of research software infrastructures
- Develop and keep official guidelines for research software management up-to-date
- Learn from and get involved in community actions
- Participate in national and international software engineering initiatives as a community service
- Contribute to and advocate open-source software
- Contribute to open science

Competence
Center
Research Data

de-RSE

DFG

Carpentries

SURESOFT

Open
Science AG

German
Reproducibility
Network

University
library

...

Example: de-RSE position paper, Research Software Directory

Initiative of the de-RSE e.V. in writing a position paper about establishing Research Software Engineering infrastructures in Germany

RESEARCH SOFTWARE DEVELOPED AT HEIDELBERG UNIVERSITY

RESEARCH SOFTWARE DIRECTORY

There is a lot of research software that is being developed and used at Heidelberg University. This Research Software Directory is meant to be a comprehensive but by no means complete list of such software, and may aid you in identifying collaboration partners.

Software Field of Study Research Group

Software	Field of Study	Research Group	DOI
A-SLOTH <small>The semi-analytical model A-SLOTH (Ancient Stars and Local Observables by Tracing Halos) is the first public code that connects the formation of the first stars and galaxies to observables. The model is based on dark matter merger trees that can either be generated based on Extended Press-Schechter theory or that can be imported from dark matter simulations.</small>	Physics, Astrophysics	Klessen group	10.21105/joss.04417 10.3847/1538-4357/ac7150
ACRONYM	Physics, Astrophysics	Spanier group	10.1007/978-3-

Support the visibility of research software by establishing a “Research Software Directory”, listing software packages developed by researchers at the University of Heidelberg

Challenges: Interdomain communication

I want to find xyz in my data.

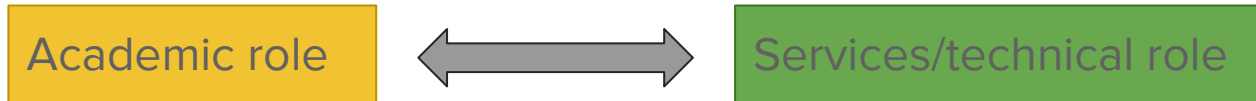
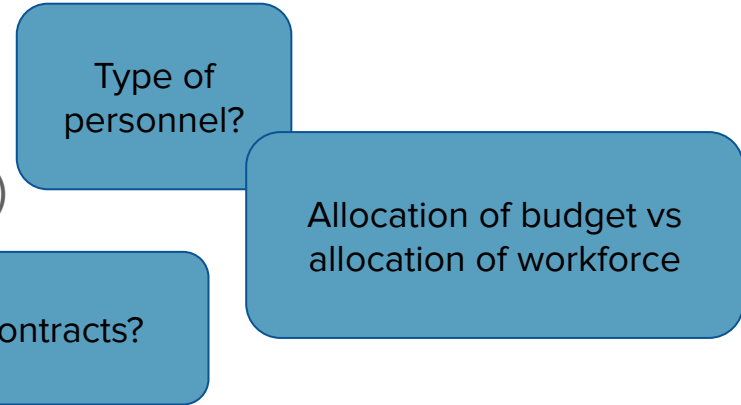
What are your requirements?



Challenges: Third-party funding

Funding options

- Directly through grant application (SSC as co-PI)
- Indirectly as service partner
- Buy-in of SSC workforce



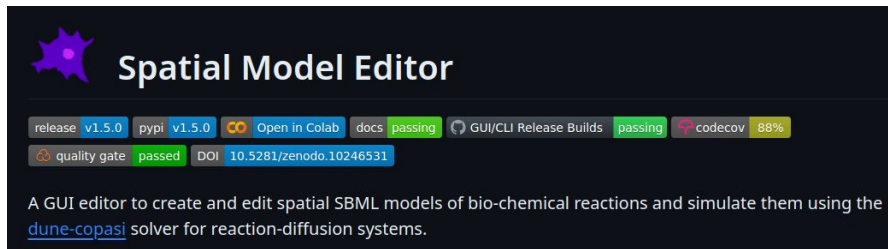
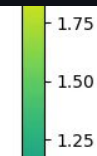
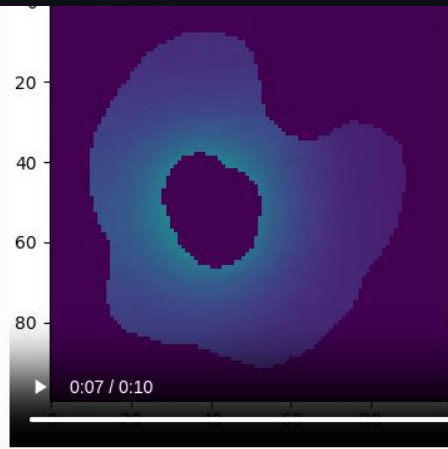
Where does an RSE fit in?

Challenges: Failures?

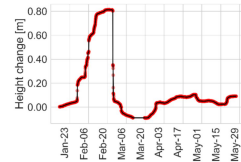
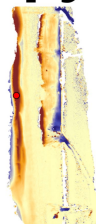
- Insufficient interdisciplinary communication
- Insufficient testing of prepared software: time constraints on the side of the researchers, failure of the research project
- Underlying dependencies not well maintained, dependency conflicts
- Data of insufficient quality
- No reference data to gauge accuracy/power of prediction

The development of research software underlies the scientific process and is subject to its failures.

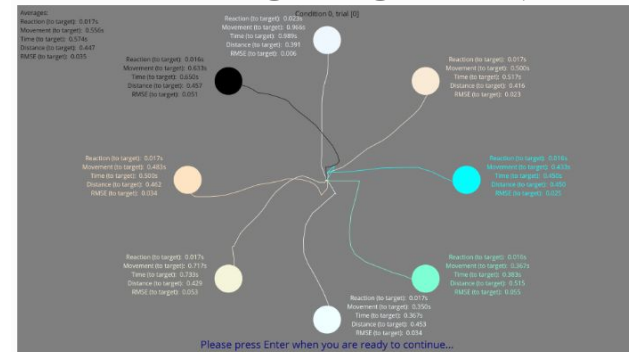
Challenges: Success stories

py4dgeo



4D objects-by-change extraction



An example of a results display after a block of trials during an experiment.

SampleFlow

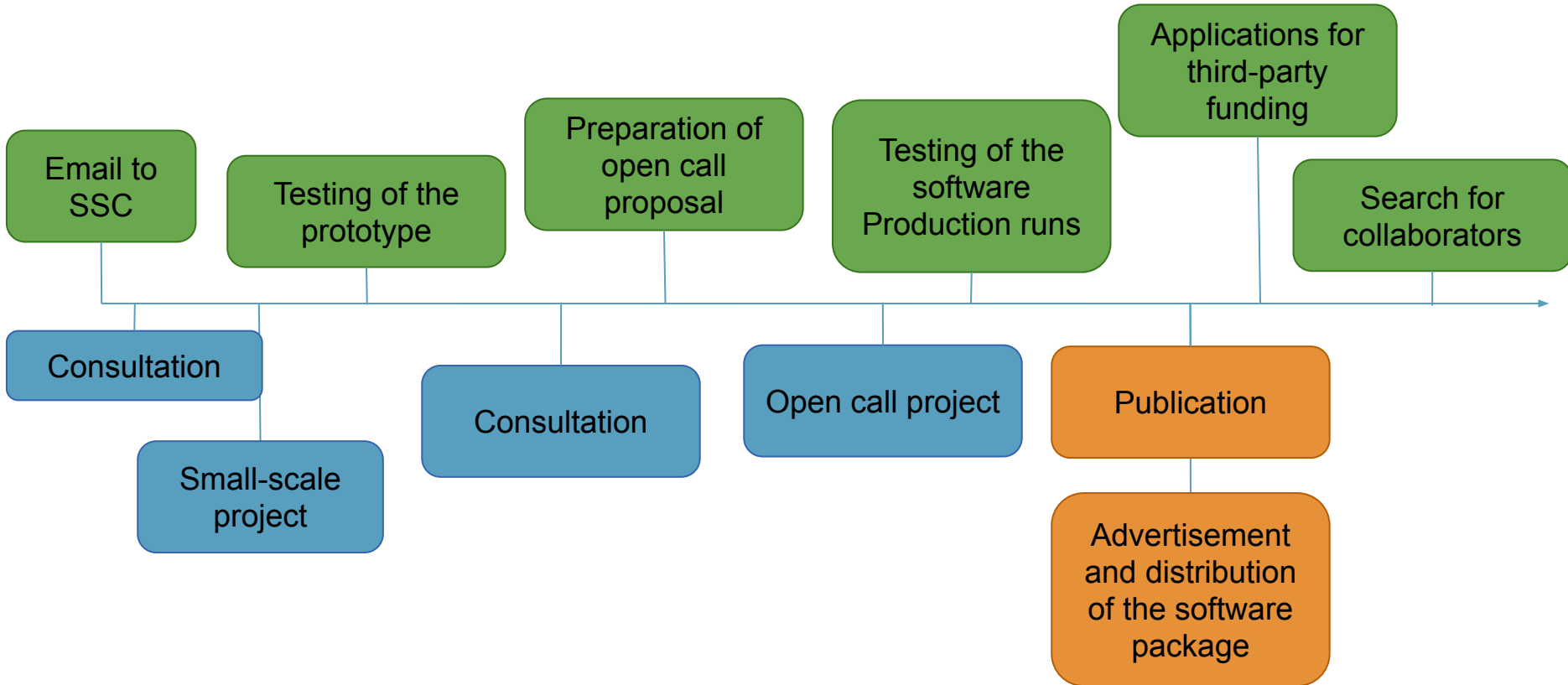
Multiplexed nanopore DNA sequencing.

[About](#) | [My samples](#) | [Login](#)

AMMICO, an AI Media and Misinformation Content Analysis Tool:
Communication research meets software development



A successful project timeline



The SSC: Want to know more?

Contact us at

ssc@iwr.uni-heidelberg.de

Take a look at our offers at

<https://www.ssc.uni-heidelberg.de>