



**CSCS**

Centro Svizzero di Calcolo Scientifico  
Swiss National Supercomputing Centre

**ETH** zürich



# Technology and Policy Requirements for Emerging Workflows: An Infrastructure Perspective

Sadaf Alam, CSCS

March 27, 2018

SOS-22

# Outline

- Motivating use cases
  - Several as-a-service requirements for community platform developers
  - Increasing diversity: data sources, workflows, access controls, etc.
- Requirements for convergence and interoperability
  - Qualitative and quantitative expectations
  - Site local and federated architecture—complementary and competing requirements
- Interactive Computing E-Infrastructure (ICEI) project
  - Federation of 5 European HPC data centres
  - Diverse technical and policy landscapes

# Research Infrastructure Driven Initiatives

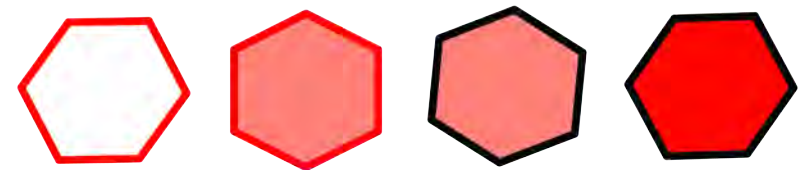
- Create and operate a European scientific **Research Infrastructure** for brain research, cognitive neuroscience, and other brain-inspired sciences
- Gather, organise and disseminate **data** describing the brain and its diseases
- **Simulate** the brain
- ...
- *Joint platforms*
  - *Collaboratory*
  - *Neuroinformatic platform*
  - ...

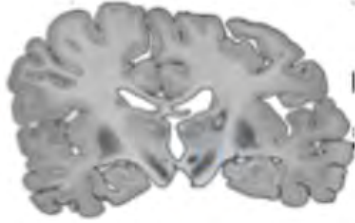


Human Brain Project

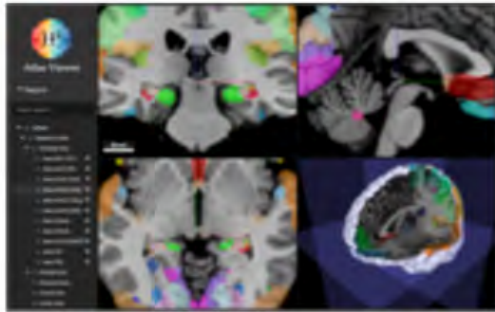
- Accelerated design and discovery of novel materials, via a materials' **informatics platform** of database-driven high-throughput quantum **simulations**, powered by
  - advanced electronic-structure capabilities
  - innovative sampling methods
  - application of **big-data** concepts to computational materials science

MARVEL

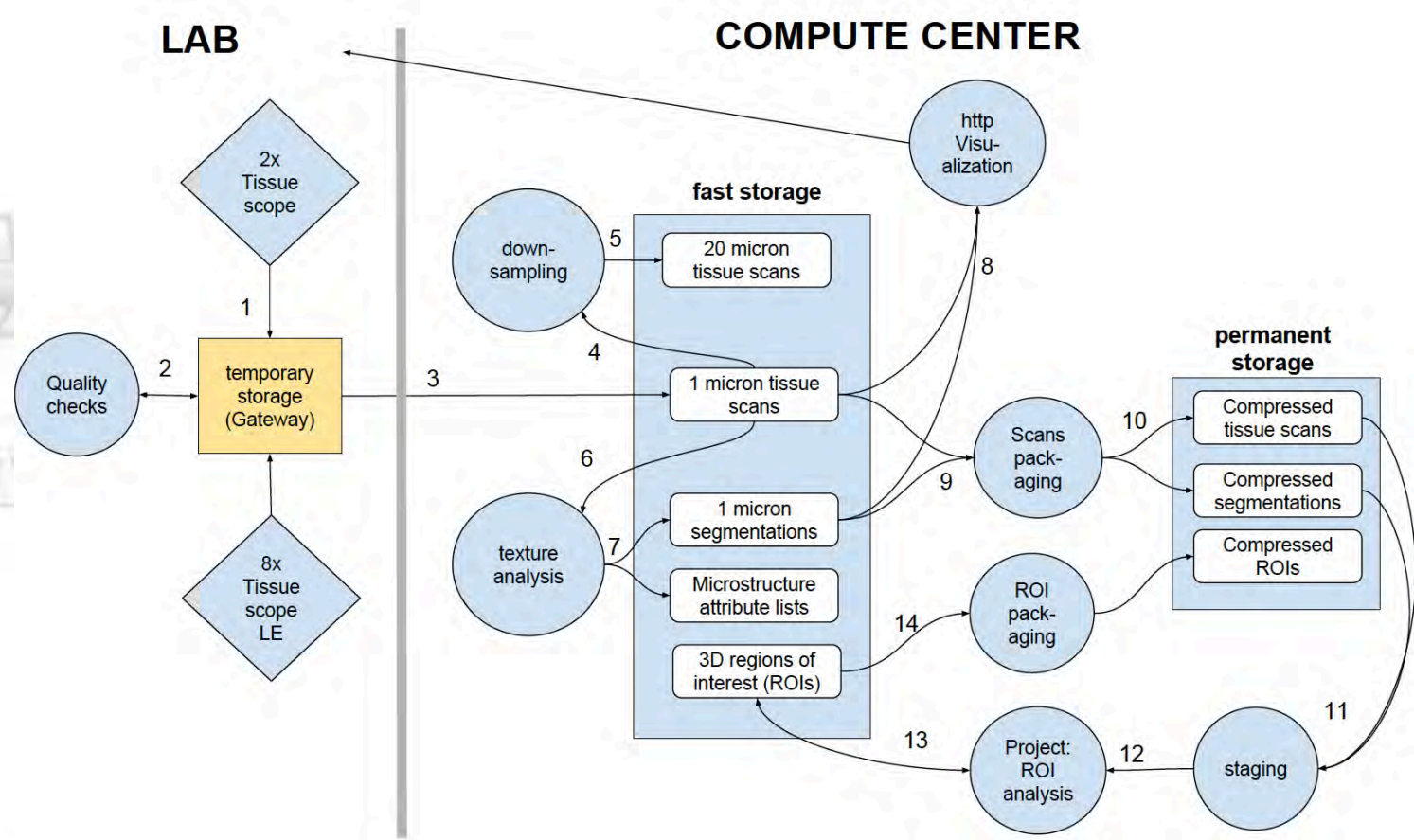




High throughput  
Microscopic  
imaging



Digital  
brain  
atlases

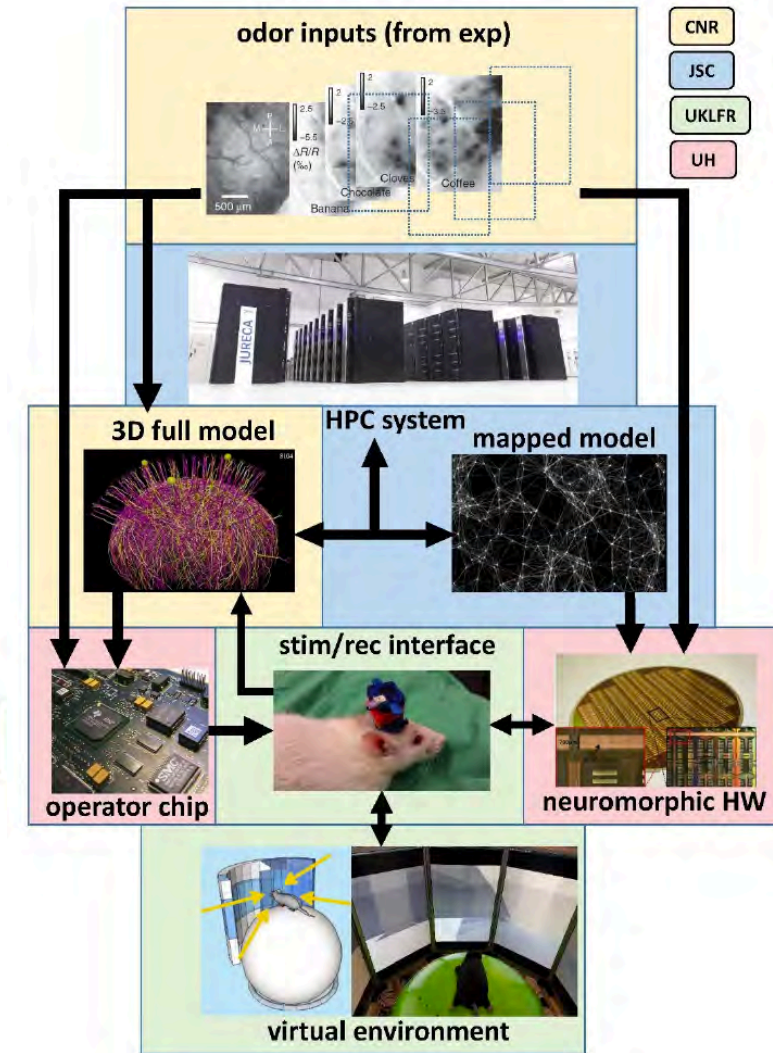
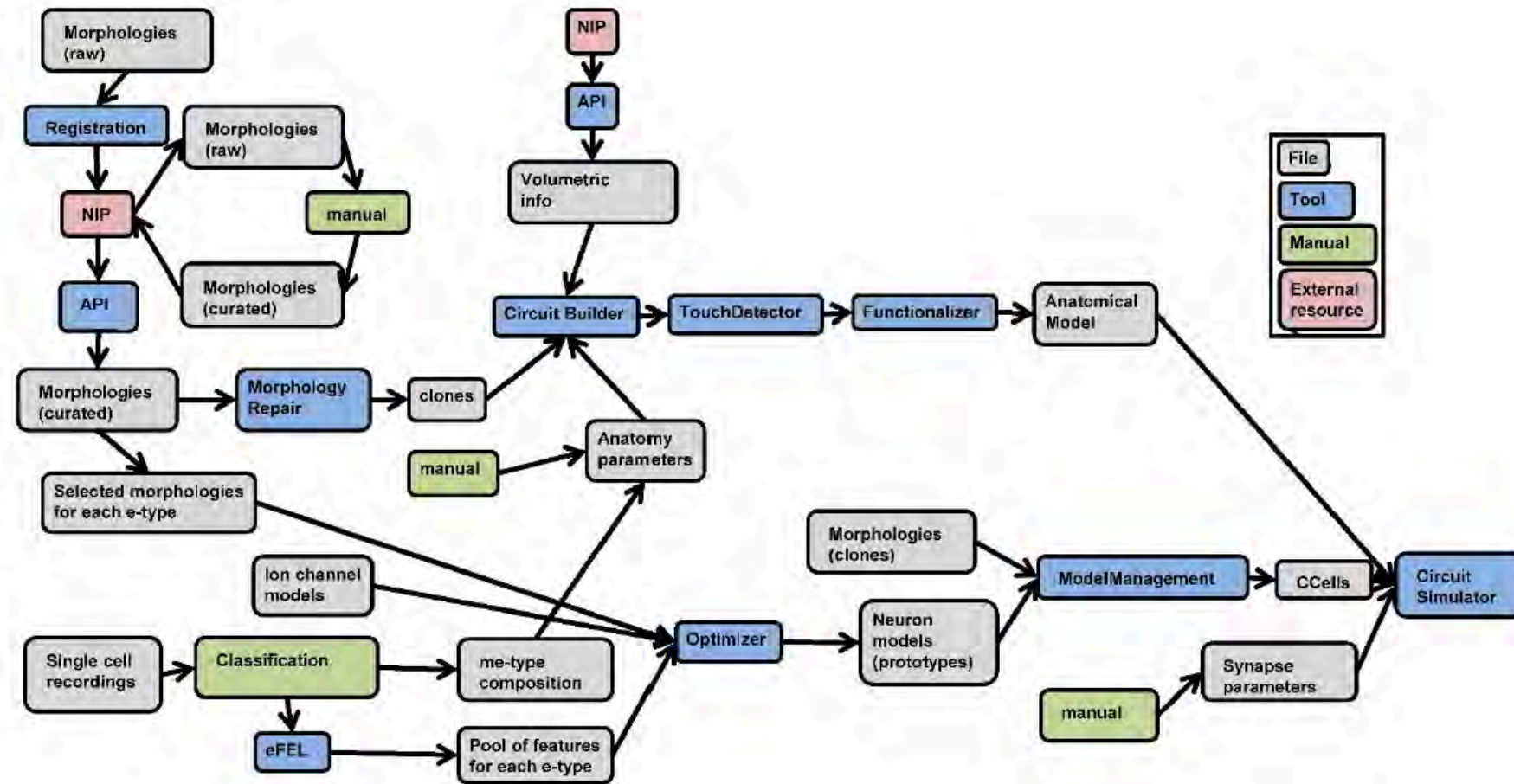


Data acquisition and analysis in the context of human brain atlasing

Timo Dickscheid, JUELICH

Co-Design Workshop on Interactive Supercomputing, February 09, 2018



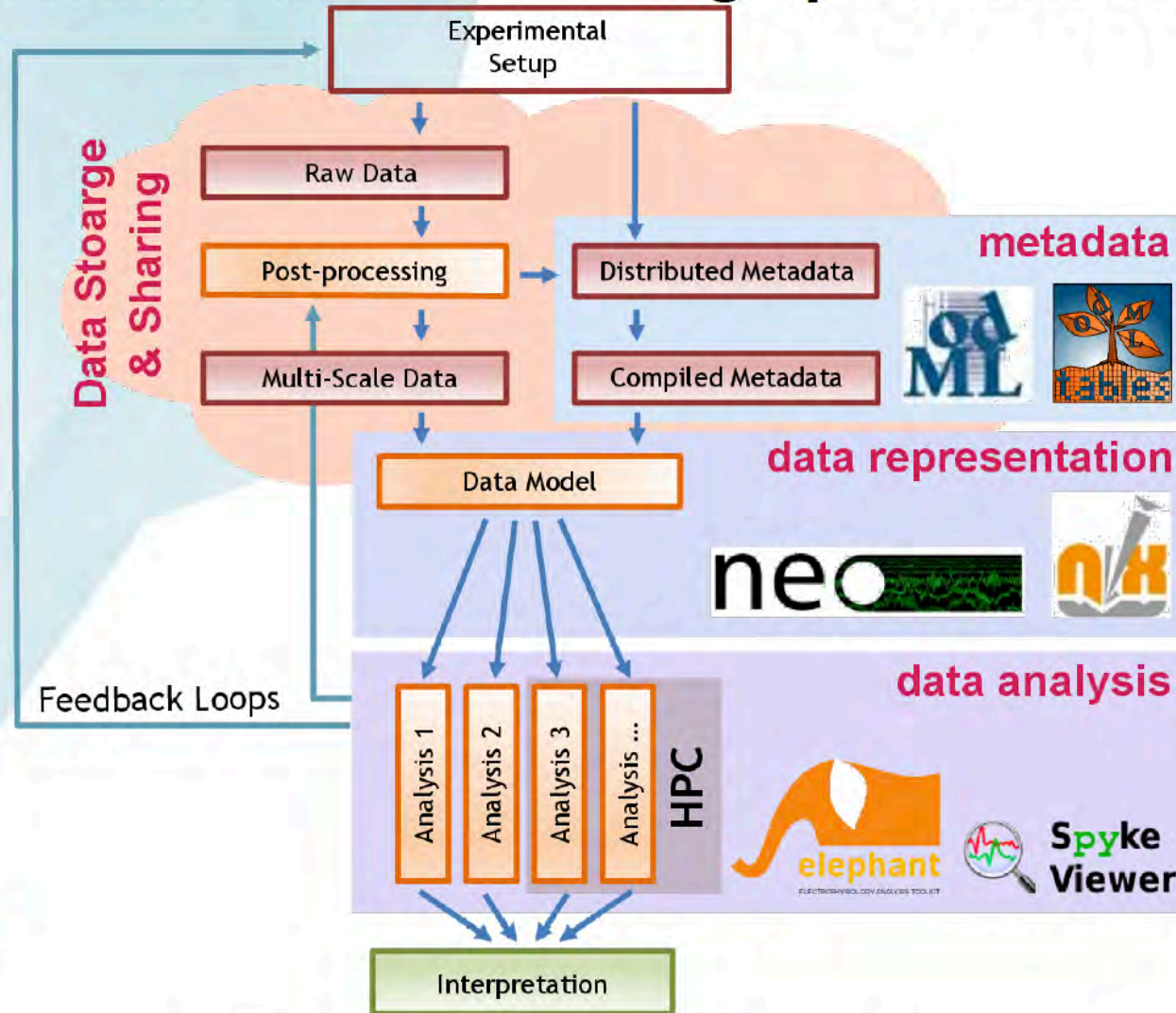


Data-driven cellular models of brain regions: the Hippocampus and the Olfactory Bulb use cases  
 Prof. Michele Migliore, Institute of Biophysics, National Research Council, Palermo, Italy  
 Co-Design Workshop on Interactive Supercomputing, February 09, 2018

# Reproducible, collaborative workflow using open software tools



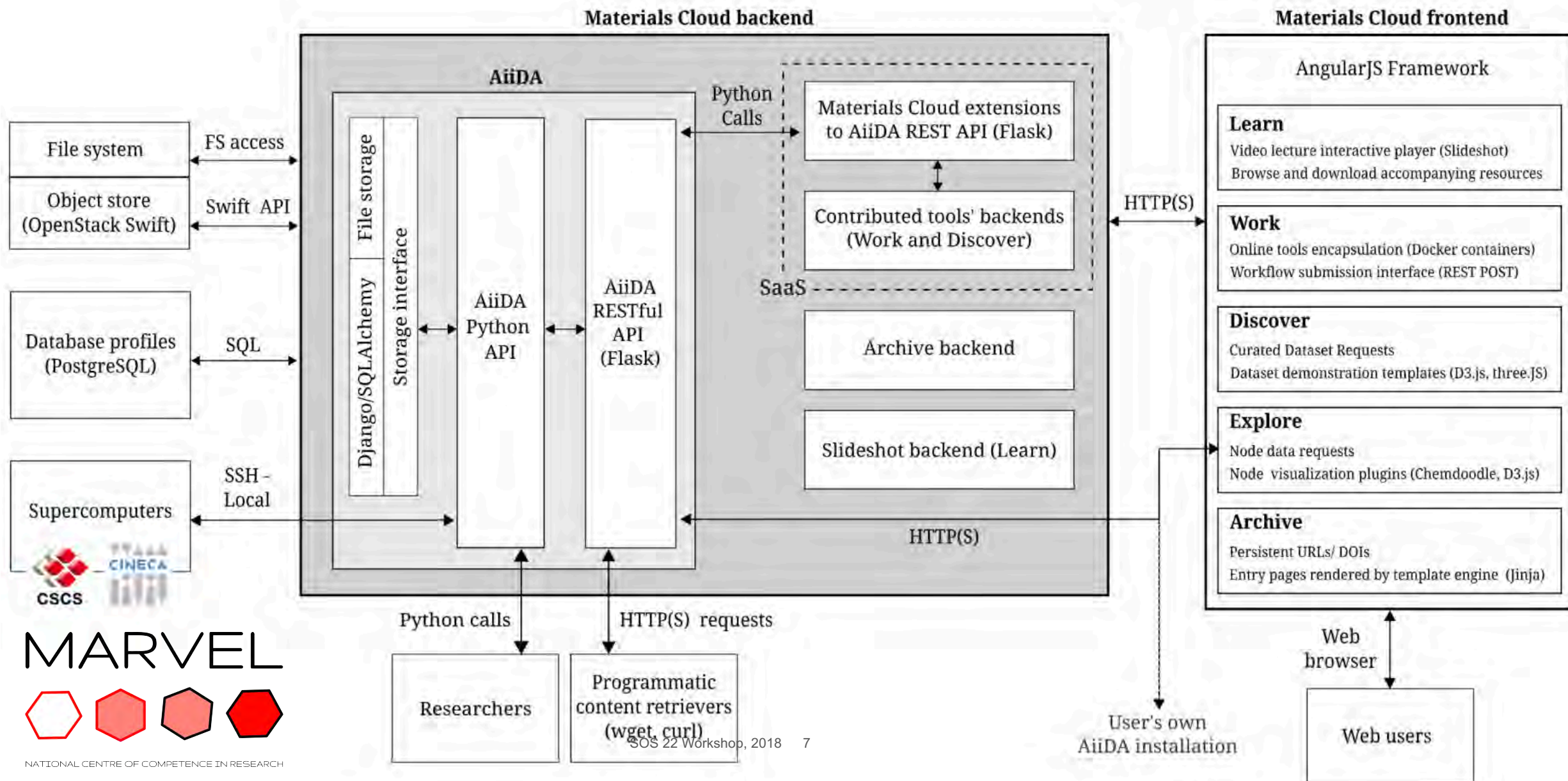
Integrated Solution



Denker, M., and Grün, S. (2016). Designing Workflows for the Reproducible Analysis of Electrophysiological Data. In Brain-Inspired Computing, K. Amunts, L. Grandinetti, T. Lippert, and N. Petkov, eds. (Cham: Springer International Publishing), pp. 58–72.



# A Platform for Open Materials Science





### Selected Profile: 2D Structures

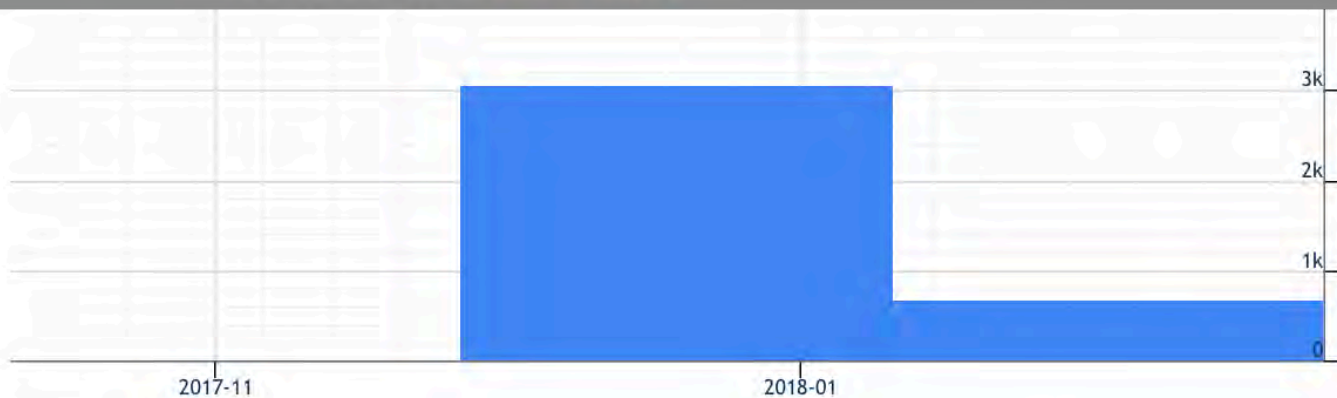
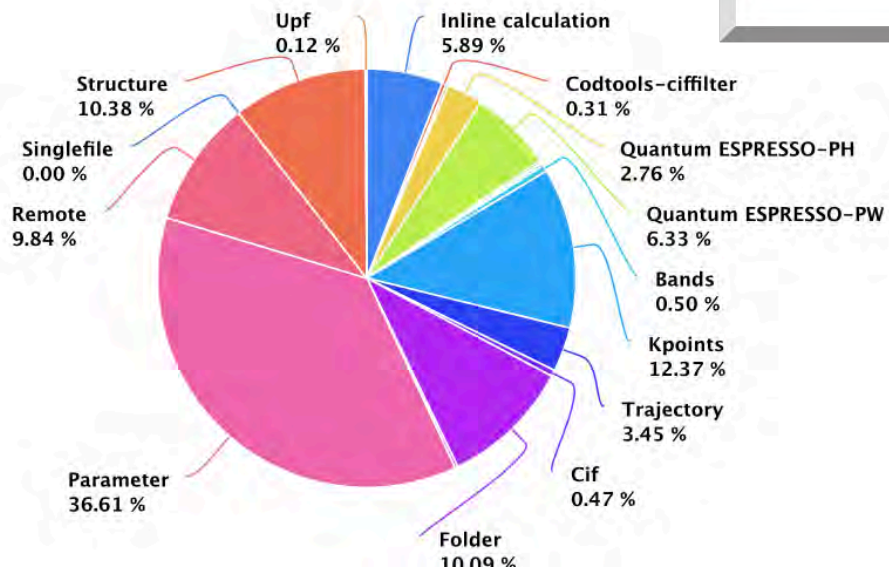
## Volume 13 Issue 3, March 2018



### Computational quest for 2D materials

Two-dimensional materials with a unique set of physical and chemical properties and the tangible potential for various electronic and optoelectronic applications have generated a substantial amount of experimental studies. Yet, all these works are based on only a few dozens of practically exfoliable materials. Using high-throughput calculations, N. Mounet and co-workers have now succeeded in identifying 1,825 potentially exfoliable... [show more](#)

Number of nodes created per type by All

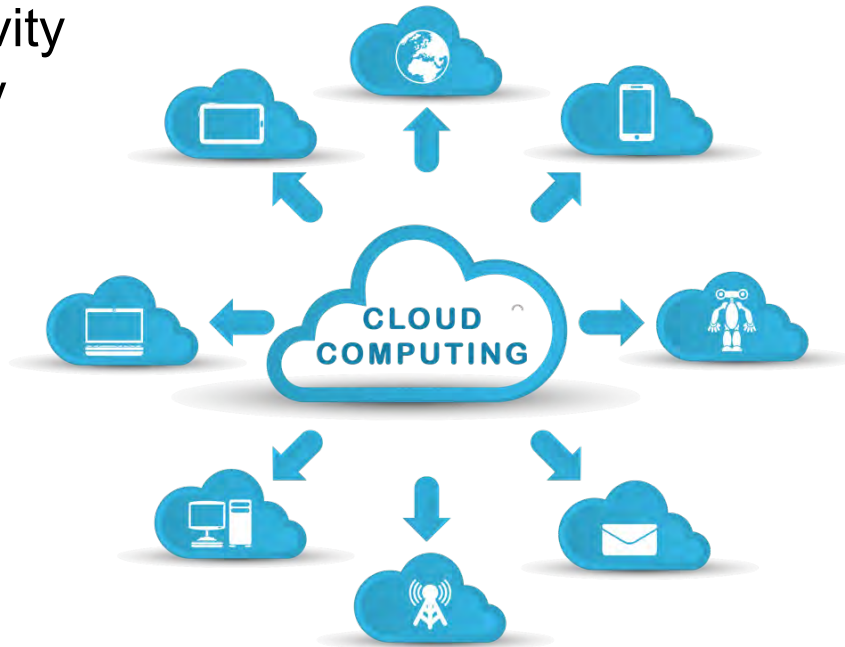




# Qualitative and Quantitative Expectations Platform Designs for Complex Workflows

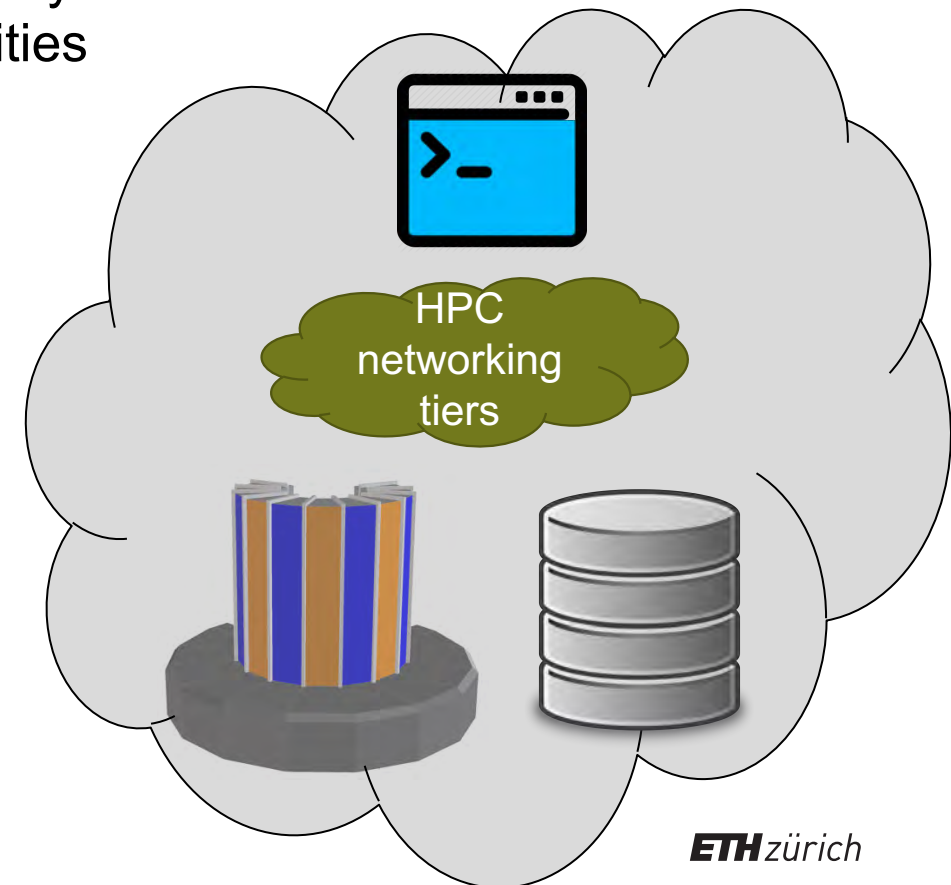
## ■ Qualitative $\leftrightarrow$ Policy implications

- Availability of the platform
- Security (anonymised data) & isolation
- User managed, fine grain access control
- Interactivity
- Elasticity
- ...



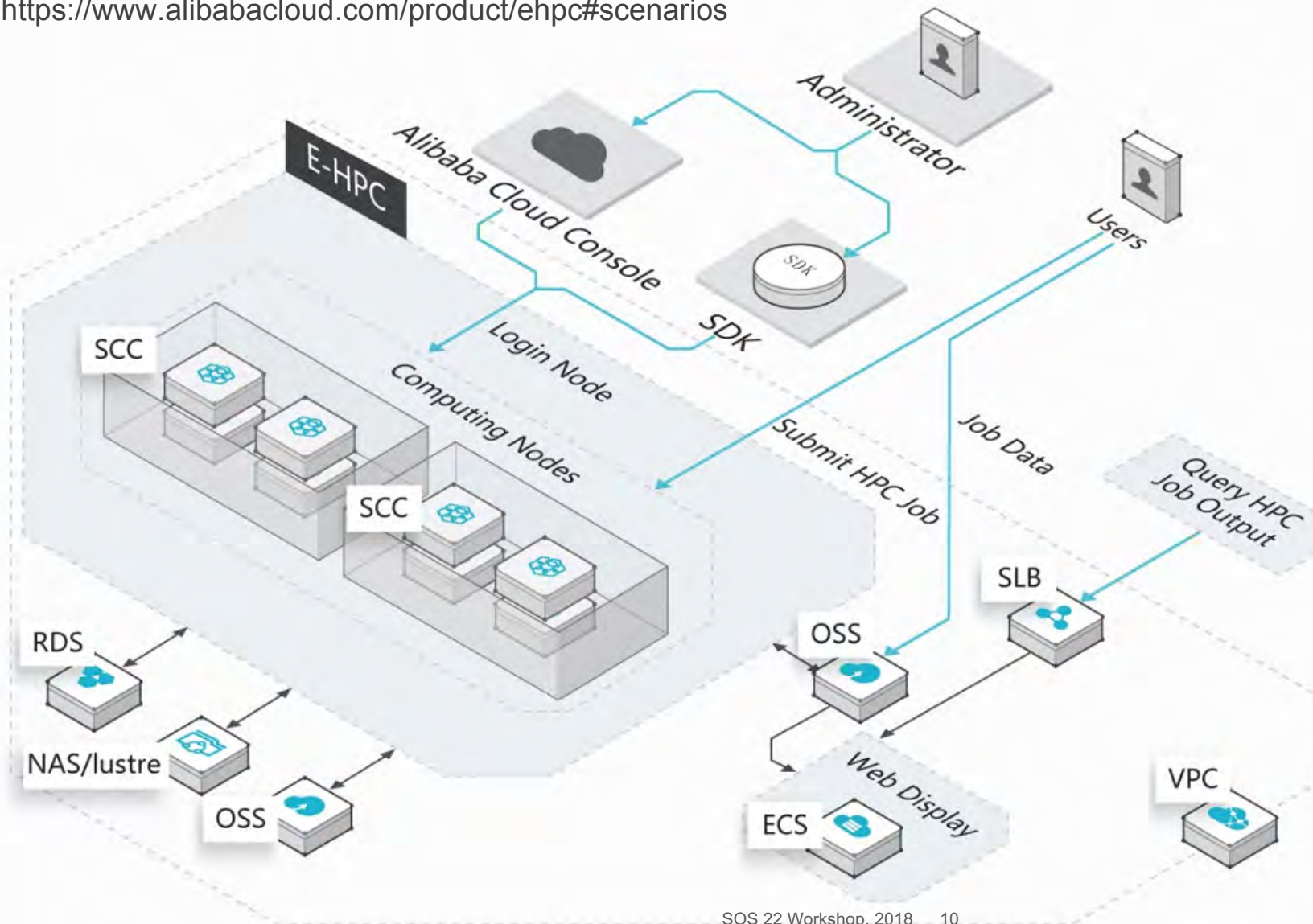
## ■ Quantitative

- Performance
- Capability
- Capacities
- ...



# Interoperability of Services (Convergence)

<https://www.alibabacloud.com/product/ehpc#scenarios>



E-HPC seamlessly integrates with other products and services of Alibaba Cloud by the user console

<https://www.alibabacloud.com/product/ehpc>

**HPCaaS**. = HPC-as-a-service  
- Infrastructure as a Service (**IaaS**) with high-performance CPU and heterogeneous computing GPU instances,  
- Platform as a Service (**PaaS**) with a high-performance computing software stack, and  
- Software as a Service (**SaaS**) with application template customization.

# Interactive Computing E-Infrastructure (ICEI)

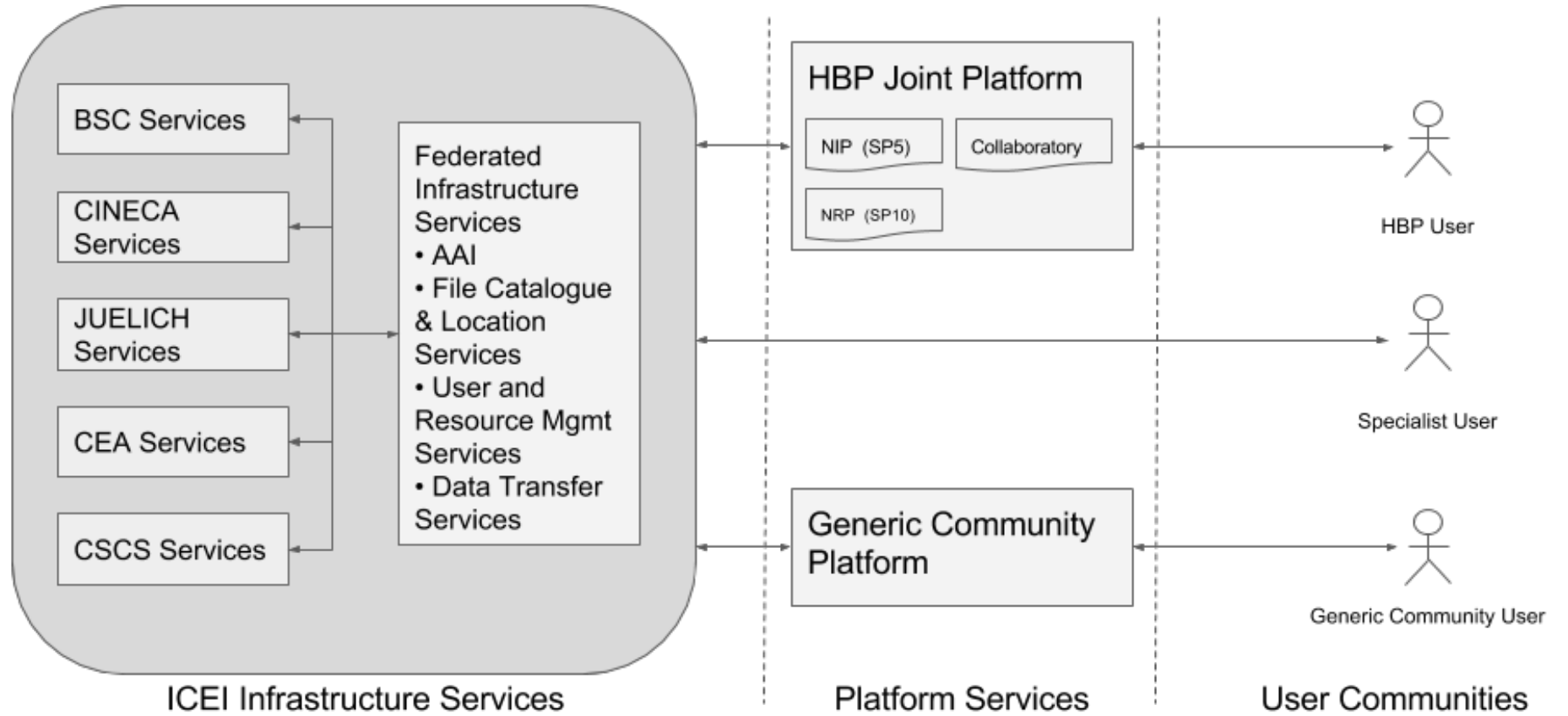


- <https://fenix-ri.eu>
- The European ICEI project is funded by the European Commission and is formed by the leading European Supercomputing Centres such as BSC (Spain), CEA (France), CINECA (Italy), ETH Zuerich/CSCS (Switzerland) and Forschungszentrum Juelich/JSC (Germany).
- The ICEI project plans to deliver a set of e-infrastructure services that will be federated to form the **Fenix Infrastructure**. The distinguishing characteristic of this e-infrastructure is that data repositories and scalable supercomputing systems will be in close proximity and well integrated.





# ICEI Overview (Site autonomy & Federated IaaS)



**Federated platform enabling services**

# Fenix Services Being Implemented by ICEI

## Technological and policy considerations

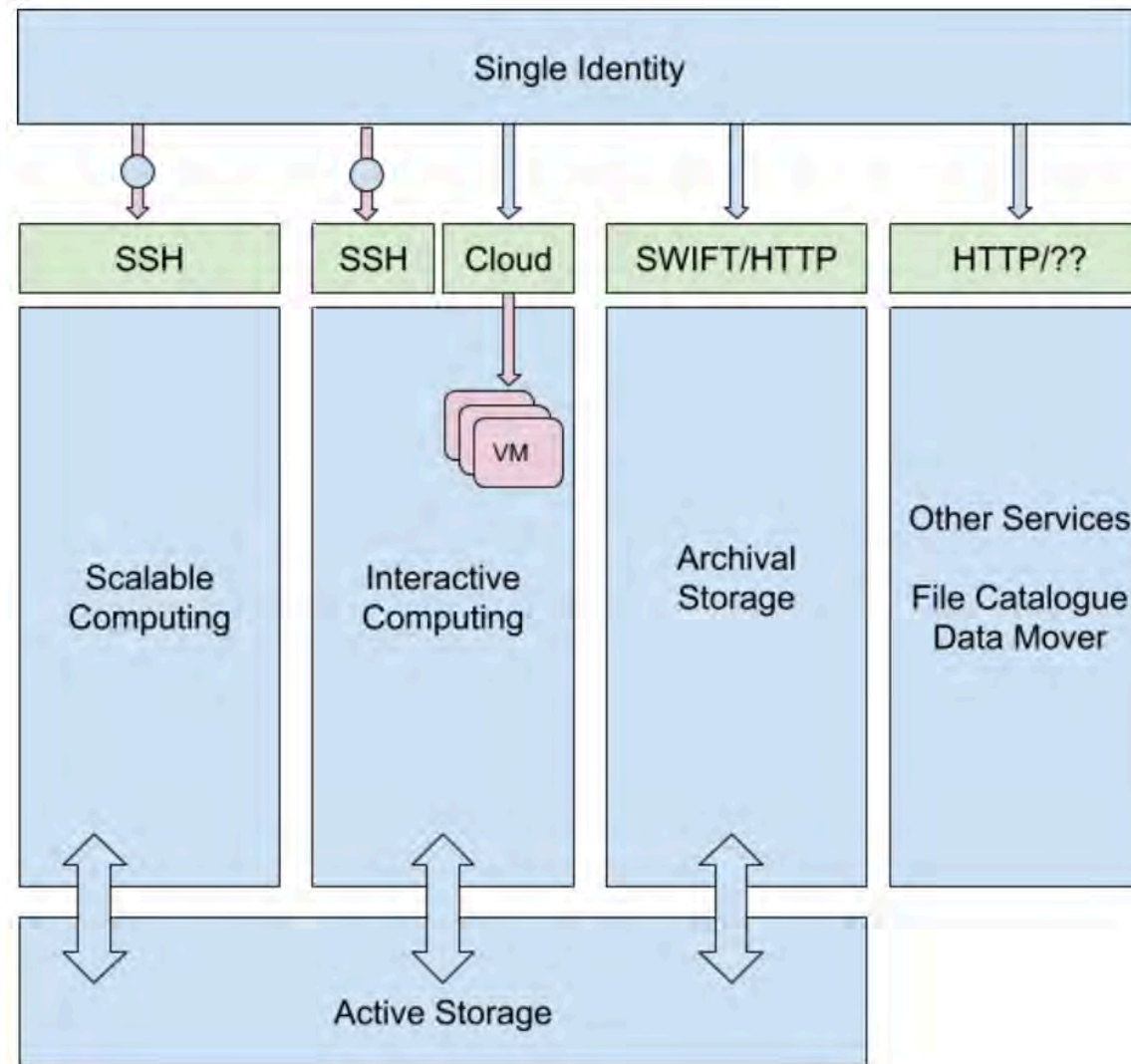
- Consumable and accountable services
  - Interactive Computing Services
  - Scalable Computing Services
  - Virtual Machine Services
  - Active Data Repositories
  - Archival Data Repositories
- Underlying and building block services
  - Internal interconnect
  - External interconnect
  - Authentication/Authorization Services
  - Data Mover Services
  - Data Transfer Services
- User and customer support services
  - Fenix User and Resource Management Service (FURMS)
  - Monitoring Services
  - User Support Services



# Envisioned Federated AAI Concept

## Technology and policy considerations

Consumable and Accountable Services
Interactive Computing Services
Scalable Computing Services
Virtual Machine Services
Active Data Repositories
Archival Data Repositories



AAI protocols:  
OpenID, OAuth 2  
and SAML

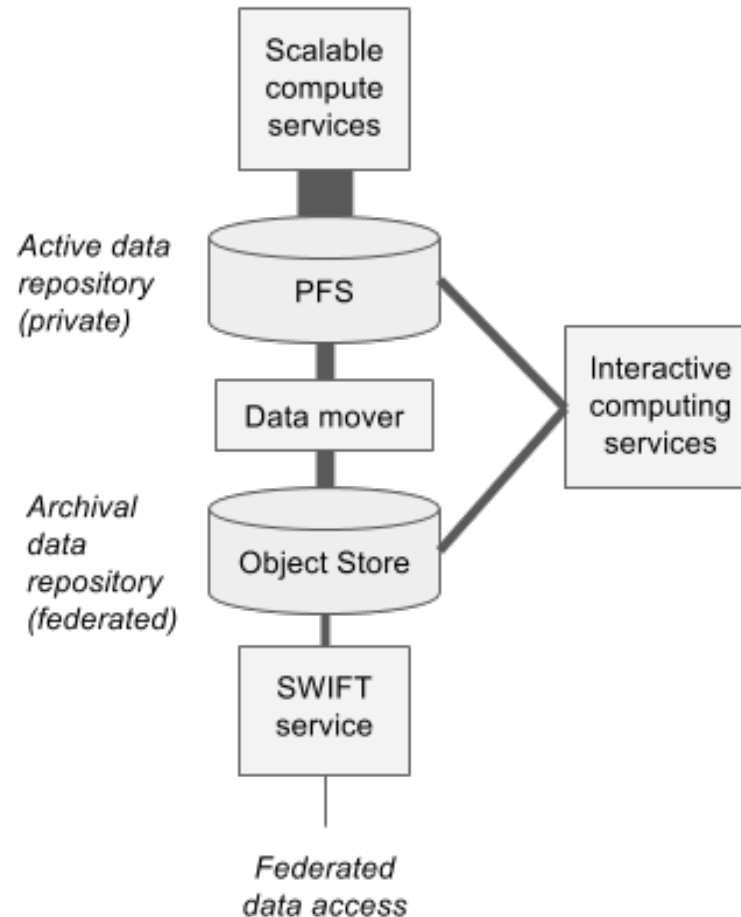
Enabling Services
Internal interconnect
External interconnect
Authentication/Authorization Services
Data Mover Services
Data Transfer Services



# Envisioned Storage Architecture

## Technology and policy considerations

Consumable and Accountable Services
Interactive Computing Services
Scalable Computing Services
Virtual Machine Services
Active Data Repositories
Archival Data Repositories

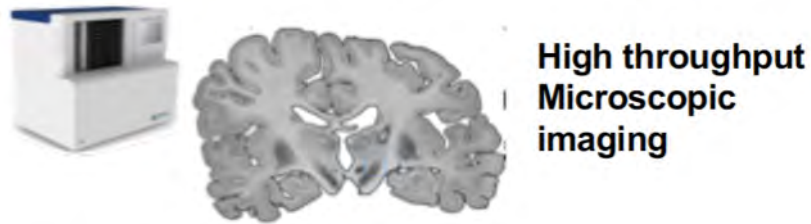


Enabling Services
Internal interconnect
External interconnect
Authentication/Authorization Services
Data Mover Services
Data Transfer Services

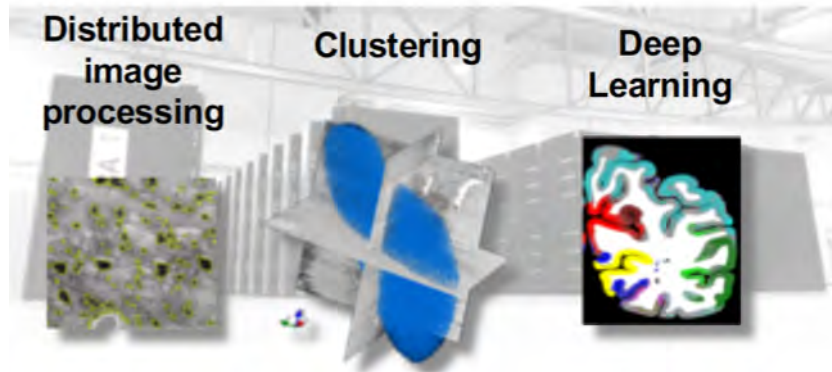
# Example of an HBP Use Case Mapping

Data sources

Neuroscience Lab

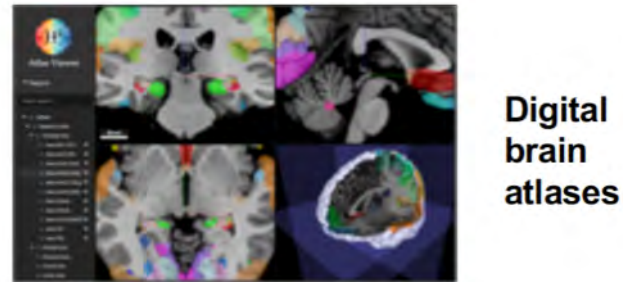


Fenix Sites



Community Platforms

User Communities



AuthN and AuthZ services
Archival data repositories
Data Mover Service
Active data repositories
Scalable computing service
Interactive computing service
Data Mover Service
Archival data repositories
Data Transport Service
Virtual Machine Services

Community platforms (highly available services)

<https://fenix-ri.eu/events/icei-public-information-event>



**CSCS**

Centro Svizzero di Calcolo Scientifico  
Swiss National Supercomputing Centre

**ETH** zürich



**Q&A**