Building infrastructures and scaling up:
Data sharing on a global scale

Session 8 | Great room 2 | 3.00-4.00PM

Hilary Hanahoe

Secretary General of the Research Data Alliance (RDA)

Dr Lesley Wyborn

Honorary Professor at the Australian National University (ANU)

Leo Chiloane

Manager of the Data Infrastructure Node (uLwazi) at the South African Environmental Observation Network (NRF-SAEON)

Nigel Smith

Executive Director and Chief Executive Officer of TRIUMF, Canada's particle accelerator centre

Inmaculada Figueroa

Deputy Director General for International Consortia, Organizations and Research Infrastructures at the Ministry of Science, Innovation, and Universities of Spain

Building infrastructures and scaling up – data sharing on a global scale

Tuesday, 3 December 2024

This session will examine the value and challenges of scaling data infrastructures from national to global levels, focusing on three key areas: economies and challenges of scale, interoperability, and the impact of AI. The panelists will discuss global system modelling with managing large data volumes, explore the potential gains and losses from standardization, and assess the costs and concerns associated with AI integration

ICRI 2024

International Conference on Research Infrastructures

3-5 December

icri2024.au

LEO CHILOANE
South African Environmental Observation Network (NRF-SAEON)

INMACULADA FIGUEROA
Ministry of Science, Innovation, and Universities of Spain

HILARY HANAHOE
Research Data Alliance (RDA)

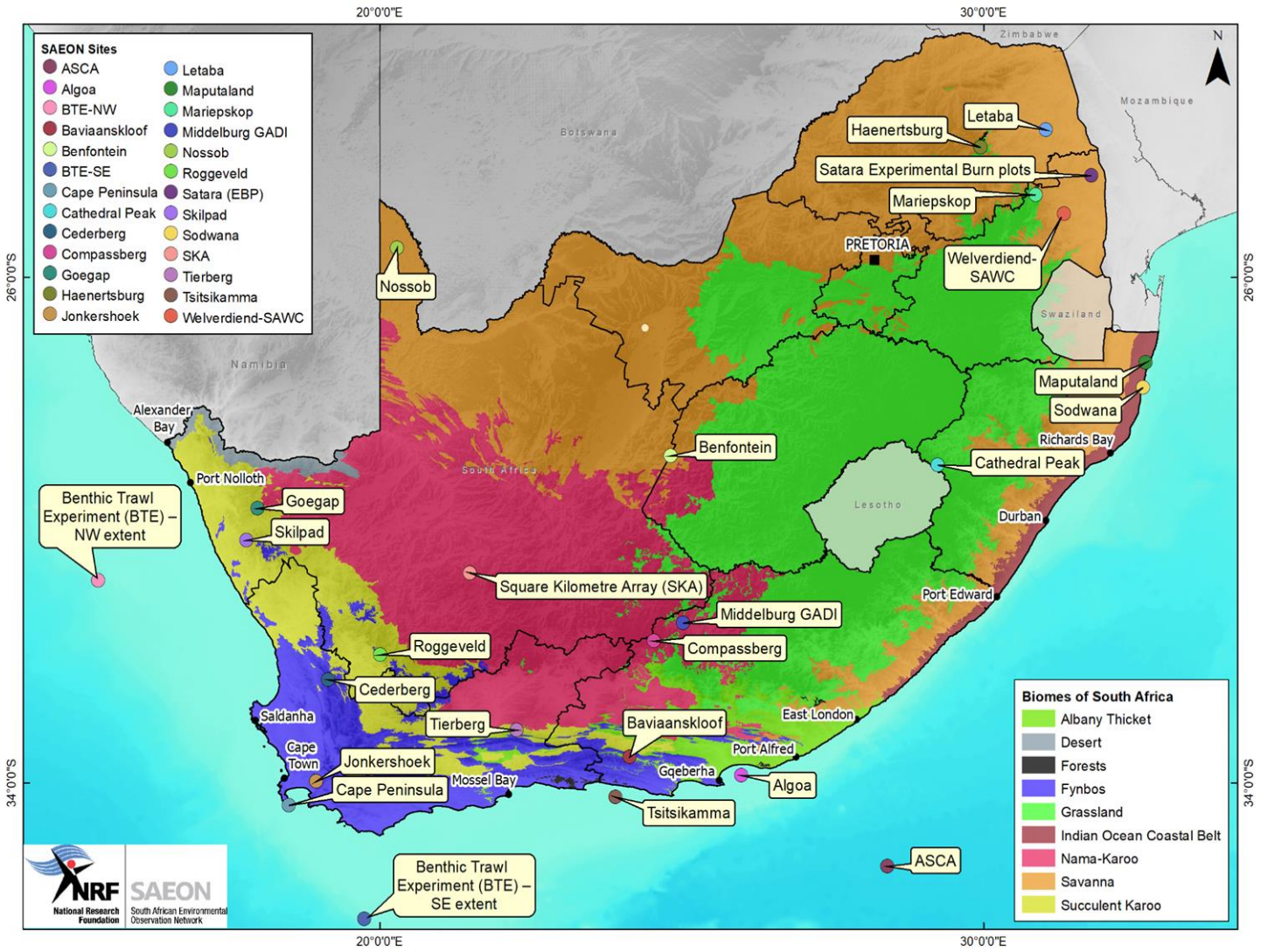
NIGEL SMITH
TRIUMF Canada's particle accelerator centre

LESLEY WYBORN
Australian National University (ANU)

PANELLIST PERSPECTIVES

icri2024.au

South African Environmental Observation Network (SAEON): An institutional data infrastructure perspective



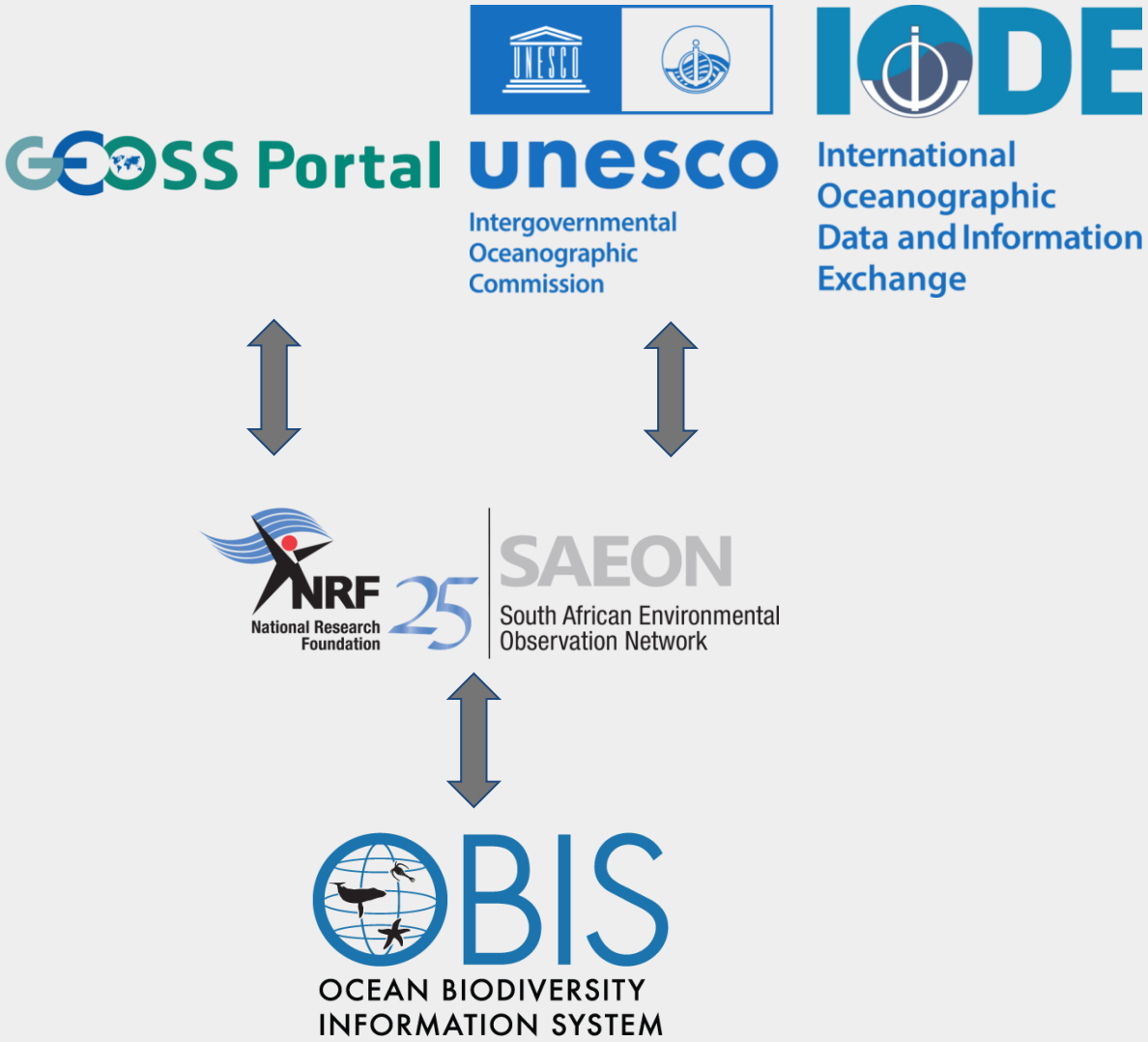
Complexities in scaling up to a global level

- **Heterogeneity:** data from multiple environmental science disciplines (marine and terrestrial) produced within the same organisation
- **Not possible to implement a universal solution** to standardisation for a multidisciplinary organisation
- **Differences in institutional data policies** a barrier to developing a centralised national and regional data infrastructure
- **Open access to data**, in a decision ready format, is essential to respond to global change

Technical implementation approach

- Developed a system of systems for data aggregation
- Development of downstream data products to enable decision support at a national level
- Focus is on sharing of metadata via standardised services to avoid cost implications of hosting large volumes of data.
- Data hosting services prioritised for partners without infrastructure

Scaling to global level, standardisation across disciplines and interoperability



Impact of AI in the organisation: collaboration and human capacity development

- Current partnership with Zindi Africa:
  - Zindi hosts the largest community of African data scientists, working to solve the world's most pressing challenges using machine learning and AI
  - They connect data scientists with organizations and provide a place to learn and connect

ZINDI

In this challenge, you must determine the precise date a flood occurred and distinguish between time series with and without flood events. By successfully predicting the day a flood occurred, the model will contribute to building a complete historical catalogue of urban floods across South Africa. This catalogue will be implemented by the South African Environmental Observation Network (SAEON), providing critical data to identify previously missed flood events and to improve future flood forecasting and response strategies.

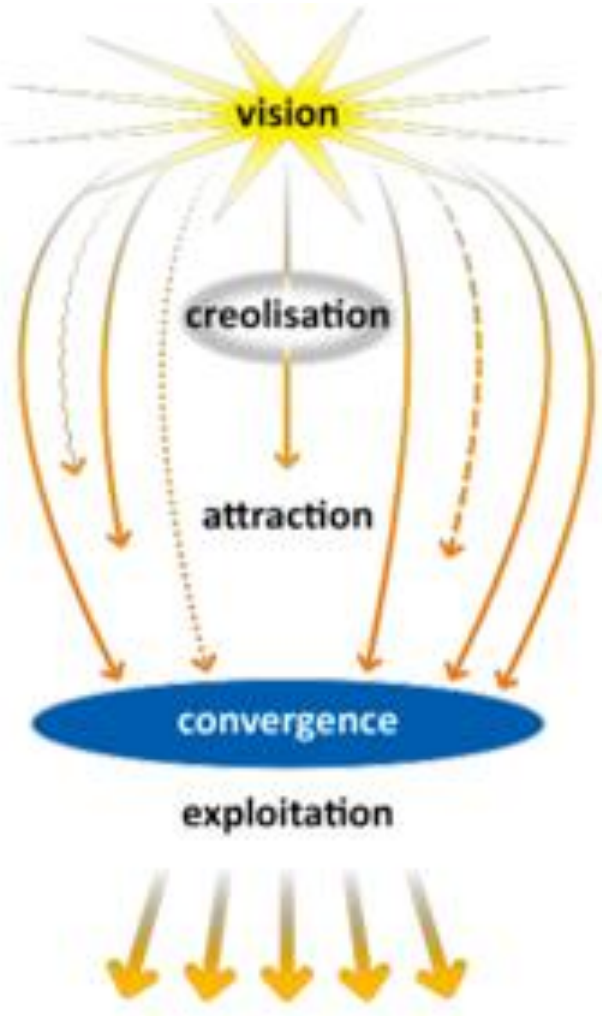
By creating a detailed archive of past urban floods, SAEON will have access to a powerful tool for analysing flood patterns, understanding the impact of urbanisation on flooding, and identifying vulnerable areas that may be at higher risk in the future. This archive can be used to improve early warning systems, optimise resource allocation during emergencies, and inform urban planning to reduce flood risk.

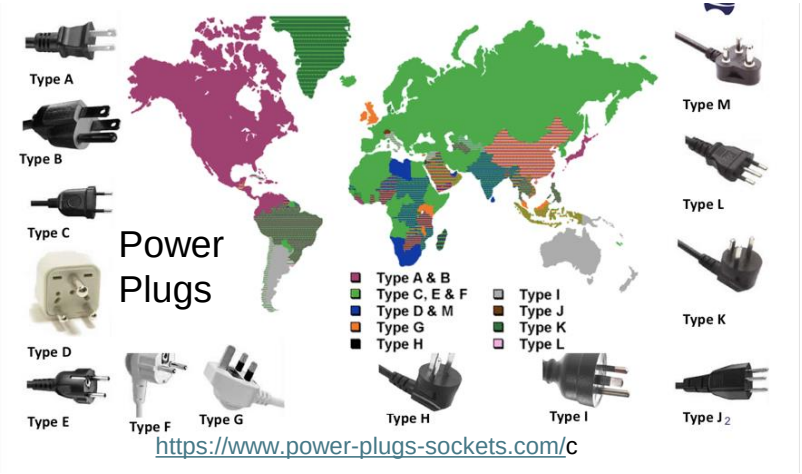
Google DeepMind

NRF 25 National Research Foundation | SAEON South African Environmental Observation Network

South African Risk and Vulnerability Atlas
MAPPING THE WAY TO A RESILIENT FUTURE

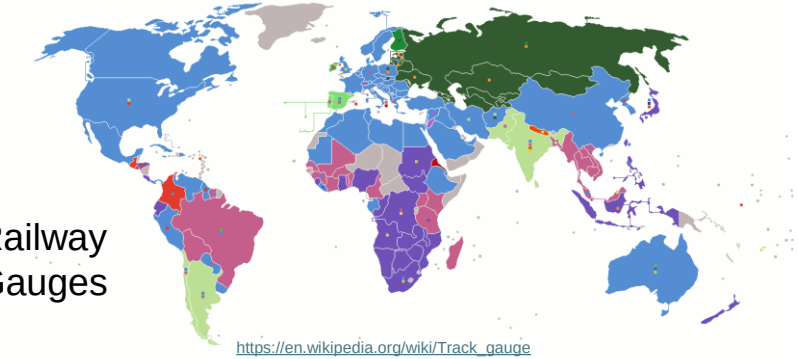
Learning from History in Building Infras





Power Plugs

https://www.power-plugs-sockets.com/c



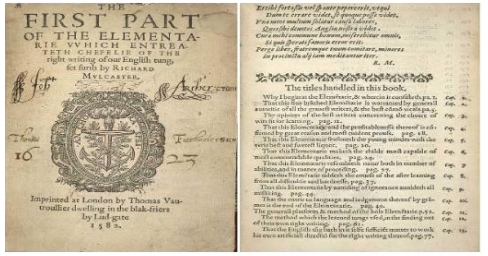
Railway Gauges

https://en.wikipedia.org/wiki/Track_gauge

mm	1676	1668	1600	1524	1520	1435	1372	1067	1050	1000	950	914	762	750	610	600
ft in	5'6"	5'5.67"	5'3"	5'	4'11.8"	4'8.5"	4'6"	3'6"	3'5.3"	3'3.4"	3'1.4"	3'	2'6"	2'5.5"	2'	1'11.6"

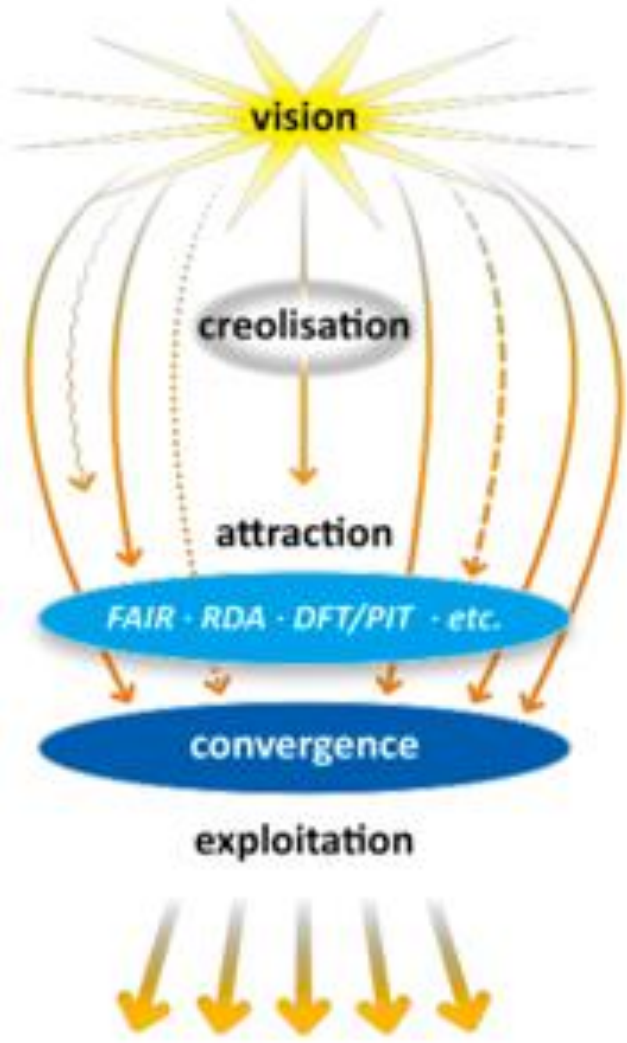
1436 AD

https://www.ripleys.com/weird-news/gutenberg-printing-press



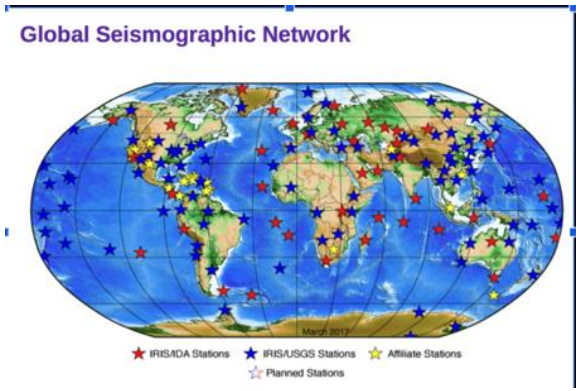
https://www.bl.uk/learning/timeline/external/mulcaster-tl.jpg

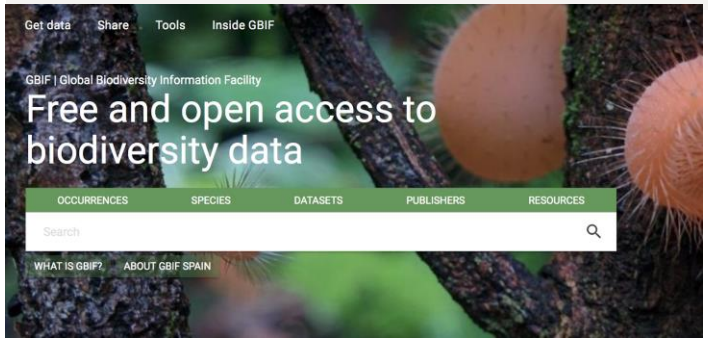
1604 AD
1755 AD

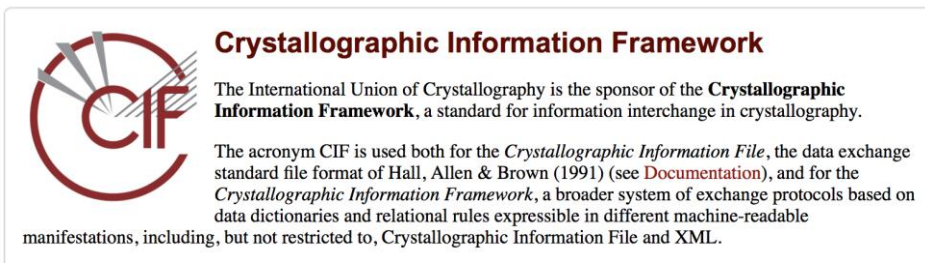


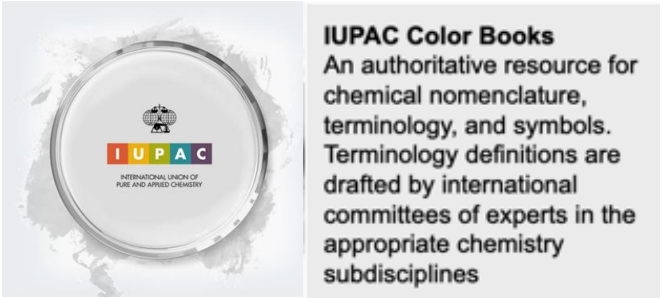
Wittenburg, P., and Strawn,G., 2018. Common Patterns in Revolutionary Infrastructures and Data
https://doi.org/10.23728/b2share.4e8ac36c0dd343da81fd9e83e72805a0

Scaling up: PC to HPC, Long Tail to Exascale – we need those dictionaries!!



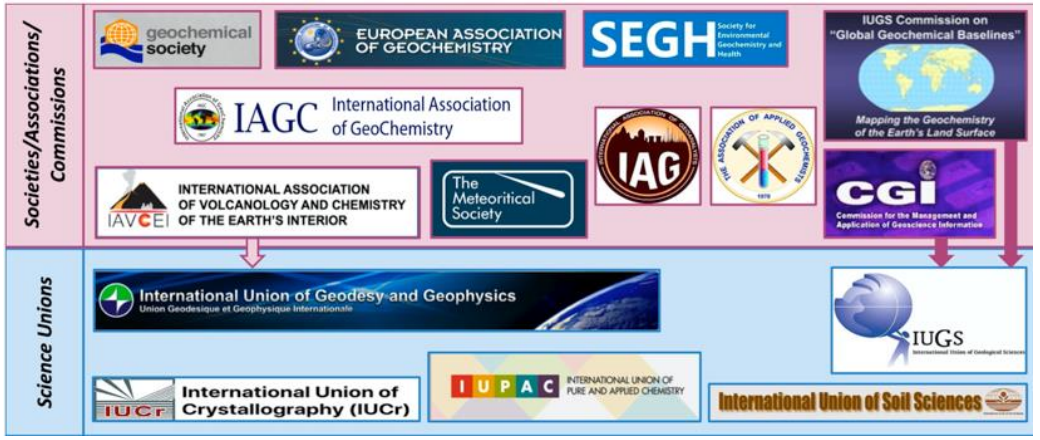






All above the line have achieved some forms of global standards for discipline scientific content

Geophysics



Geochemistry

What is TRIUMF?

TRIUMF is Canada's particle accelerator centre. We are a world-class hub of research, education, and innovation that is home to ~600 staff and students, utilising a suite of particle accelerators, including the world's largest

TRIUMF operates at the interface between academia, government, and industry. Owned and operated by a consortium of 21 Canadian universities

Supported by federal and provincial governments, the Canadian research funding agencies and the Canada Foundation for Innovation, TRIUMF hosts ~$2B of infrastructure and $100M annual operational budget

TRIUMF’s Research Programme

TRIUMF’s work **addresses the most compelling challenges in contemporary science** and connects fundamental scientific research through to commercialisation.

We act as an interface to major sub-atomic physics communities and research infrastructures around the world, such as CERN and J-PARC

TRIUMF Innovations translates the work we do for societal benefit, such as the development of new therapeutic radioisotopes and electronics irradiation services

TRIUMF is located on the traditional, ancestral, and unceded territory of the Musqueam people, who for millennia have passed on their culture, history, and traditions from one generation to the next on this site.

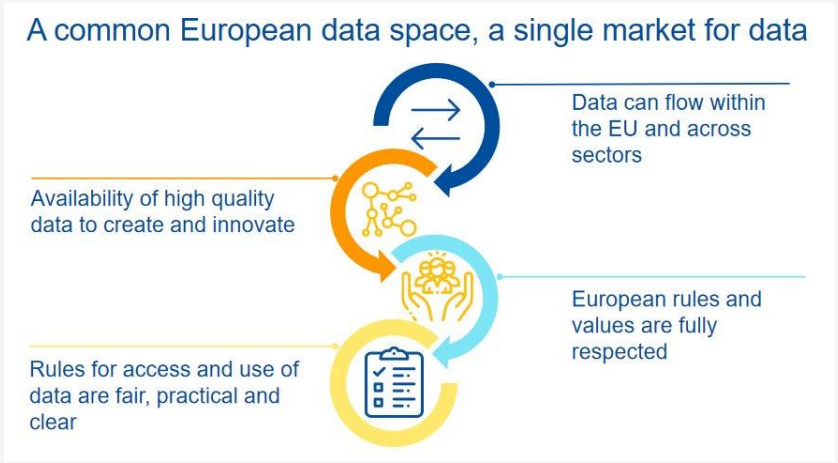
TRIUMF’s home on the UBC campus has always been a seat of learning.

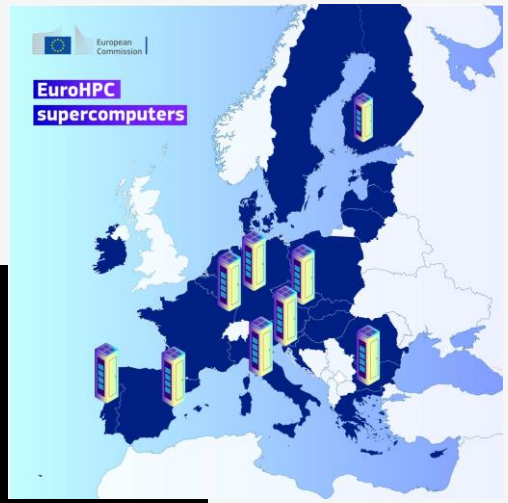
eosc

ESFRI

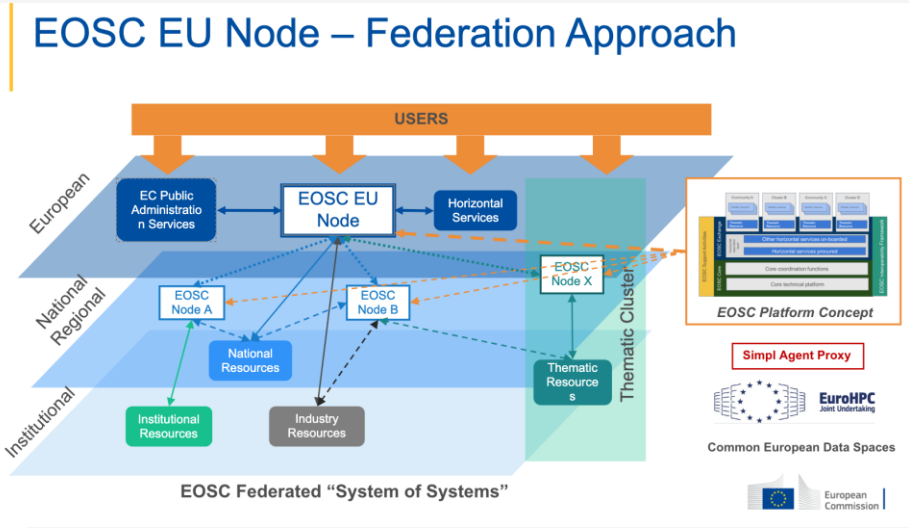
ESFRI-EOSC Task Force

EuroHPC Joint Undertaking





- LUMI
- LEONARDO
- MareNostrum5
- MeluXina
- Vega
- Discoverer
- Karolina
- Deucalion
- JUPITER in Germany
- SOON Jules Verne and Daedalus



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

PANEL DISCUSSION

With respect to economies of scale, international or cross border or cross disciplinary cooperation and collaboration both supports and augments opportunities and investments for research infrastructures. Would you agree with that statement?

When speaking with each of you, people and skills emerge as an important component of successful infrastructures. Is there sufficient emphasis on and funding for the human component of data infrastructures?

With respect to data interoperability, we have the FAIR (Findable, Accessible, Interoperable and Reusable) principles and now the CARE (Collective benefits, Authority to control, Responsibility and Ethics) principles, couldn't infrastructures just implement them (data standards) and the problem would be solved?

Artificial Intelligence (AI): is here to stay and there are many positive and negative statements concerning AI. With respect to data and research infrastructures will it resolve all the challenges that still have to be tackled or will it complicate things even more?

Do you see AI, the availability and lack of resources and competencies, creating a divide across the globe? Creating a two-tier system where the “haves” define the processes, standards, protocols, etc. while the “have nots” just have to take those on board? Within the context of data and research infrastructures what are your top 2-3 urgent things that need to be done now to avoid such a scenario

CLOSING STATEMENTS

Save the Date!

INTERNATIONAL DATA WEEK

13-16 October 2025
Brisbane, Australia

HOSTED BY

ARDC Australian Research Data Commons

Australian Government

NCRIS National Research Infrastructure for Australia

ARDC is enabled by NCRIS

CODATA Committee on Data International Science Council

RDA Research Data Alliance

World Data System

OUR PARTNERS

Business Events Australia

Tourism & Events Queensland

brisbane australia

Brisbane Convention & Exhibition Centre

idw2025.org

THANK YOU