

Europe joins forces to create largest ever shared data repository for researchers

15 June 2017

by Benedict O'Donnell

Research institutes and policymakers have agreed to create the European Open Science Cloud to provide access to data from all publicly funded research carried out in the EU. Image credit: CC0

World-leading research institutes have agreed to join forces with funding agencies and policymakers to create the European Open Science Cloud, the largest shared data repository in history.

The idea is to give every scientific user access to the data resulting from research carried out with public funding, using a single login.

According to Professor Iain Mattaj, head of the European Molecular Biology Laboratory in Heidelberg, Germany, the cloud-based service could facilitate the sort of interdisciplinary collaboration that has transformed our present understanding of the earth's climate and the human genome.

By 2020, it could grant the science community direct and easy access to scientific data generated by publicly funded research, and used for scientific publication.

Broad agreement on the go-ahead came during the European Open Science Cloud Summit in Brussels, Belgium, on 12 June.

Carlos Moedas, the European Commissioner for Research, Science and Innovation, described the European Open Science Cloud as a one-stop shop for scientists to pool their observations, and as a commons for European science.

'I see it as a new Republic of Letters,' he said, referring to the correspondence between intellectuals during the Age of Enlightenment.

The platform could save time on unnecessary duplication of experiments and grant even small research groups access to vast data samples.

'Science has outgrown the constraints of technology once again. It needs this internet of data to continue.'

Carlos Moedas, European Commissioner for Research, Science and Innovation

This unprecedented pooling of research data from across scientific disciplines and national borders takes to a new level the open research data repositories that have supported scientific collaboration since the early days of the internet.

No one is underestimating the scale of the challenges to be faced in making the European Open Science Cloud work.

'We are talking about a staggering amount of data,' said Dr Harald Reichert from the European Synchrotron Radiation Facility in Grenoble, France.

His institute alone records petabytes (millions of gigabytes) of measurements each month – enough to drown even Facebook's servers over just a couple of years.

Making this volume of research data accessible to 1.7 million researchers across Europe requires vast amounts of bandwidth, funding and work, he said.

Professor John Womersley from the European Spallation Source super microscope in Lund, Sweden, insisted that storing data is not enough. The success of the European Open Science Cloud will hinge on how easy it is to access and use.

'The internet would be useless if it was just data,' said Prof. Womersley. 'Researchers also need tools like browsers, search engines and apps that find the data that they are looking for and turn it into something that they can use.'

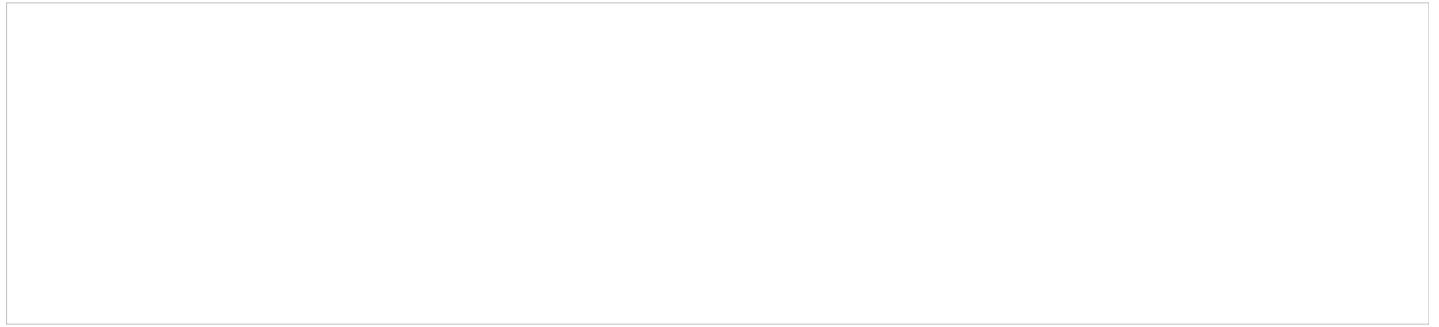
Decades

This raises unprecedented data management questions, said Professor Eckhard Elsen from the CERN (European Organization for Nuclear Research) particle accelerator in Geneva, Switzerland. 'In research fields like climatology, data need to be averaged over decades before they can be used,' he said. 'A private company can go bankrupt over those periods.'

To avert the risk of the science cloud becoming a research data graveyard, these tools have to meet the needs of research communities across different sciences, with each dataset annotated and formatted in a way that one single system can make sense of.

According to Dr Reichert, annotating the details of each dataset would be too time-consuming for scientists to do themselves – all the more so in cross-disciplinary research, where the requirements of peers are not necessarily known.

Research data 'stewards' would be needed, in addition to strong incentives and skills for all scientists to be able to share research data, in order to ensure the quality of the data.



Volunteers are now formalising their commitments in a joint declaration, and aim to share tasks in a structured, open strategy, with the agreement of EU Member States.

The national science funds of the Netherlands and Switzerland have agreed to provide advice on making the data released by new research projects internationally accessible, and other participants have offered their experience in data exchange platforms spanning borders and specialisations.

Commissioner Moedas announced that the EU itself would invest over EUR 200 million over the next two years into coordinating on-going efforts.

'We want to put our money where our mouth is and get the European Open Science Cloud running by 2020,' said Commissioner Moedas. 'Science has outgrown the constraints of technology once again. It needs this internet of research data to continue.'

If you liked this article, please consider sharing it on social media.

Next steps

The idea is for the open science cloud summit to result in a formal declaration, which will be agreed upon by all stakeholders.

There will be a plan on how to achieve the European Open Science Cloud by the end of this year, according to the European Commission. This will include a strategy for governance and funding.

A dedicated governance board will also handle the implementation of the open science cloud by 2020.

More info

[European Open Science Cloud](#)