

Digital preservation: Alliance set to tackle science's new frontier

November 26 2007

A new digital divide, or rather chasm, is opening up in the scientific enterprise, and something urgently needs to be done to prevent data from being lost into oblivion. At the Second International Conference on Permanent Access to the Records of Science held in Brussels on the 15th November, the Alliance for Permanent Access, a group of stakeholders dedicated to preserving digital science records, was launched to do just that.

“We are addressing a very serious problem in maintaining accessibility to the work of scientists and what they have done in past generations,” said Peter Tindemans, acting chair of the Alliance, and President of Global Knowledge Strategies & Partnership, “This requires collaborative efforts of key stakeholders in the research enterprise”.

The Alliance for Permanent Access brings together major international and national scientific organisations such as European Science Foundation (ESF), CERN, ESA, Max Planck Society and libraries which have joined forces to help create a European digital information infrastructure.

The Conference brought together 60 experts and representatives of partners in the Alliance for Permanent Access to the Records of Science, to discuss how the preservation of digital science publications and data can be embedded into scientific practice across Europe.

Why keep the data?

“The first email was sent in 1964,” said Lucy Nowell of the United States National Scientific Foundation, “but that first email has been lost forever.” This historic moment went the way of the 13,000 NASA tape recordings of the first mission to the moon. Since the 1960s vast amounts of digital data, now measured in petabytes (one quadrillion bytes), equivalent to a kilometre-high stack of CDs, has been produced through increasingly complex experiments often taking place on a global scale. The question is can the world afford to lose this data”

There is no doubt that implementing preservation strategies will be costly, although how much investment is required is still an unknown. In general stakeholders agree that data must be preserved in a way that guarantees open access, interoperability so that datasets can be compared within and across scientific fields, and repositories must be developed to meet these needs in a quality-controlled and sustainable manner. On the flip-side the unknown cost of losing data makes evaluating preservation more difficult still.

With the first beams planned to circle CERN’s 27 kilometre Large Hadron Collider (LHC) in May 2008, the issue of storage is more than urgent. When it is fully operational the LHC experiment, which aims to recreate conditions a fraction of a second after the big bang, will be generating 15 petabytes of information per year.

Experiments like these produce data that cannot be replicated and require storage solutions to preserve data in a useable form for future generations to analyse, compare and re-use, yet as Jos Engelen, Deputy Director General of CERN, admitted, “We do not have a real long-term archival strategy to access this data.”

“From the point of view of a high-energy physicist, scientific data is

complicated because preserving our data in a digestible form that doesn't require details such as exactly how the experiment was carried out and the weather conditions on the day, is difficult", said Engelen.

Wouter Los, an ecologist from the Hungarian Academy of Sciences, explained another aspect to data conservation in the analysis of interlocking systems: "Using pre-existing data allows us to create and analyze scenarios and probabilities to understand how diseases and parasites are introduced into Europe. This is a totally new approach," added Los, "so we need to ensure that the scientists can easily use all these kinds of data, and that the data is interoperable."

A change of culture

What is needed is a change of culture, something which the European Union has already recognised. Focusing on digitisation and digital preservation, the European Commission is taking on the role of leveraging stakeholders and developing policy initiatives on a strategic and technical level. Though projects tend to take a broad view some science-specific work is underway. In February of this year the Commission issued a Communication on "Scientific information in the digital age" and is promoting discussion via high level and member state groups. The Commission is also taking a market-based approach to establishing the economic incentives to preserving data, with a proposal underway to develop a study on the socio-economic drivers and impact of longer-term digital preservation.

A European Digital Information Infrastructure

Along with the EU, the Alliance has committed to spreading good practices and to promoting research and development into preservation and management tools. With the goal of creating a European Digital

Information Infrastructure, the Alliance has identified scientific communities as the key structural approach to meeting the challenges ahead. In addition it will focus on developing funding models and economic analyses to assess the cost of sharing and accessing data and identify ways in which these costs can be integrated into all funding mechanisms for science.

The next steps for the Alliance include the creation of a forum on preservation and access and developing a handbook of good practices. The Alliance is also hoping to secure funding to develop tools from available European Union programmes.

“The initiative is courageous because there are so many people, communities, views involved but it is going to be a challenge to develop something sustainable and useful. I think the acting chairman, Peter Tindemans, is very energetic, he has the right vision, but now he has to secure the right sort of collaboration. And the patronage of the ESF is crucial,” concluded Engelen.

ource: European Science Foundation

Citation: Digital preservation: Alliance set to tackle science's new frontier (2007, November 26) retrieved 23 April 2024 from <https://phys.org/news/2007-11-digital-alliance-tackle-science-frontier.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.