# Dealing with data: Upgrading infrastructure

Los, W.; Wood, J.

DOI
10.1126/science.331.6024.1515-c
Publication date
2011
Document Version
Submitted manuscript
Published in
Science

**Pre-print**

# Dealing with data: Upgrading infrastructure

Published in: W. Los, J. Wood in: Science, Vol. 331, (2011), p. 1515-1516

The SPECIAL SECTION on DEALING WITH DATA (11 February 2011, p. 692) illustrates the emerging crisis of a data deluge. Two diagrams accompanying the introduction suggest that a radical change is required. One shows the response to the question where researchers archive most of their generated data, and 50.2 % of the respondents refer to their own lab and 38,5 % mention university servers. The question as to whether there is sufficient funding for their data curation 80.3 % is answered with a firm NO. The basic problem is that most researchers manage data in their own storage software and hardware system. They are forced to address data persistence and data curation, while the outside world is pressing them to share the data. The resulting picture is a clumsy network of poorly linked databases of unknown quality. It is a cottage industry trying to cooperate in an international environment.

At the AAAS 2011 meeting in Washington DC we presented the conclusions and recommendations of a European high level expert group on scientific data entitled "Riding the Wave – How Europe can gain from the rising tide of scientific data" *(1)*. They address the implications of an international framework for a collaborative Data Infrastructure. Rather than preserving current practices, a radical change is proposed. Dedicated Data Services should provide secure persistent storage, data identifiers, authenticity and workflow support for data mining and other applications to access data for multidisciplinary interests. Note this is not a grand unified service but probably a number of separate ones with data controlled by the generators and owners, and operated via a number of public and/or private services who accept common standards. This Data Infrastructure is not the supporting technology, but the "data world" as a reliable source for knowledge. The drive for this service will become reality when researchers accept the need for economic and efficiency reasons. In addition to adapting the technology further to make data capture in this way easier there is a need to develop the appropriate legal and financial models which will foster trust in this service. There is a need to agree the next steps on international stage to ensure common actions to the radical change.

WOUTER LOS,[1] JOHN WOOD[2]
[1]University of Amsterdam, Institute for Biodiversity and Ecosystem Dynamics, NL-1098XH Amsterdam, E-mail: w.los@uva.nl, [2] Association of Commonwealth Universities, London WC1H 9HF, UK.

## References

1. http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf