Dealing with Big Data

"'Big data' is definitely here to stay; the question is how we deal with it: how will we store and curate the data? Who will be responsible for their management? How can we ensure interoperability? On these questions, the RDA can provide important input. Digital techno-logies are not only creating more data, they also provide us the tools to make sense of them. Text and Data Mining (TDM) can analyse and extract new knowledge and enables new research connections. A recent report indicates that prolific use of TDM would add tens of billions of euros in value to the EU's aggregate GDP."

What will the 'data world' look like in 2020?

Robert-Jan Smits, Director-General of the European Commission's DG Research & Innovation   credit EC DG Research & Innovation

"Systemic changes are taking place in the way research is performed and science is organized, which are sometimes referred to as Science 2.0. In 2020 the data world will be shaped by the main drivers of Science 2.0, which include amongst others the tremendous increase in the number of researchers, new emerging scientific powerhouses, the growing and increasingly pressing demand for solutions to grand challenges, or 'digital natives' becoming part of the researcher population."

http://www.iprhelpdesk.eu/ Open_Access_in_H2020020.pdf.

Science 2.0

Science 2.0 is a term used to describe the ongoing evolution in the modus operandi of doing research and organizing science. These changes in the dynamics of science and research are enabled by digital technologies and driven by the globalisation of the scientific community as well as the increasing societal demand to address the Grand Challenges of our times.
ec.europa.eu/research/ consultations/science-2.0

COLUMN

"Let's always discuss what research must be like"

One day when my son was in his final undergraduate year at university he asked if I could tell him what it was like to do a PhD. Amazed that he was asking his father for advice I started to tell him what it was like for me and my students.

It was only afterwards that I realized I had not told him what I thought research would be like in the future. I was then prompted by a member of the European Commission to write up my thoughts on what research would look like in 20 years' time. Parts of this were subsequently published in Microsoft's Futures magazine and I thought no more about until I got a phone call from what was DG Infso to ask if I would chair a high level group on the future of scientific data. While I had led organizations that created and stored huge amounts of data from international experiments, I am not a computer or data scientist. Fortunately the members of the group were fantastic and we had tremendous support from the Commission in addition to contributions from a number of key witnesses. The result was 'Riding the Wave' or how to deal and cope with the tsunami of research data. Published at the end of 2010 it became a best seller since not only were the recommendations clear, but they resonated with other initiatives going on in other parts of the world. One of the key recommendations was to ensure there was a global forum to avoid different approaches by different regions of the world.

Further discussions took place between Europe and the USA and these culminated in a pre-meeting at the ICRI2012 meeting in Copenhagen where colleagues from other interested countries, including Australia, met to discuss future actions. Towards the end of it Alan Blatecky from the National Science Foundation stood up and said: "Let's get on with it!" After a year of intense activity the Research Data Alliance was formed between the USA, Europe and Australia. Right from the start it was agreed that the RDA was not an academic debating chamber but was about tangible results that could be used by the community. The sheer energy and momentum of the RDA has surprised all of us.

Is it a false hope that research around tackling global challenges will lead to a sustainable future?

Only time will tell but the involvement of groups looking at agriculture, marine and social issues among many gives hope that the cultural change of sharing in an Open Science and Open Innovation environment will lead to a sustainable future for the world. Tall ambitions, but essential if the earth is to sustain 9 billion people in 2050.
The original Riding the Wave report looked forward to 2030. It is likely that many of the recommendations will be in place before 2020. So it is time to look ahead even further. My granddaughter is just over a year old, what will research be like when she starts her PhD? Discuss!

john.wood@acu.ac.uk

Professor John Wood, RDA Coun-cil Co-Chair, Secretary-General of the Association of Commonwealth Universities and former High Level Expert Group on Scientific Data Information Chair & European Research Area Board Chair, mem-ber of the Research, Innovation and Science Expert Group Science 2.0.   credit Association of Com-monwealth Universities

---

Combined data services of partners

Secure research data with one click of a button

Research Data Netherlands (RDNL), founded in 2013, promotes sustainable archiving and reuse of research data. Jeroen Rombouts, working at one of the promoting organizations, tells us more about this 'open national coalition.'

For Dutch researchers it is getting increasingly easy to secure research data for future use. It may even require no more than one click of a button. That is, if it is up to the RDNL partners. Jeroen: "RDNL is a coalition of back office parties and currently made up of three partners: DANS, 3TU.Datacentrum and SURFsara. From RDNL, they offer their services to their customers as an integrated back office. If, for example, the customer is a university data library, the data library purchases data services from RDNL, which they then provide to their researchers. For staff members of research institutes this will lead to a more customer-friendly service offering and practice, particularly where cross-discipline and/or multi-supplier issues are at stake.

Maintaining research project results could be entrusted to, for example, a single front office which records the data and then stores them in different repositories under similar conditions."

Ensured data value

He continues: "The partners benefit greatly from the coalition too. They complement each other through knowledge pooling and transfer. They also look for more efficient employment of their collective capacity. Moreover, cooperation improves the position of the individual parties at the national and international levels. Their main mission is to ensure the accessibility, usability and long-term availability of valuable data. To live up to their expecta-tions, RDNL will coordinate its roadmap with various stake-holders, including users, policy makers and research funders. The partners will then address these actions both together and with stakeholders from outside RDNL. 2015 will be an important year!"
J.P.Rombouts@tudelft.nl
researchdata.nl
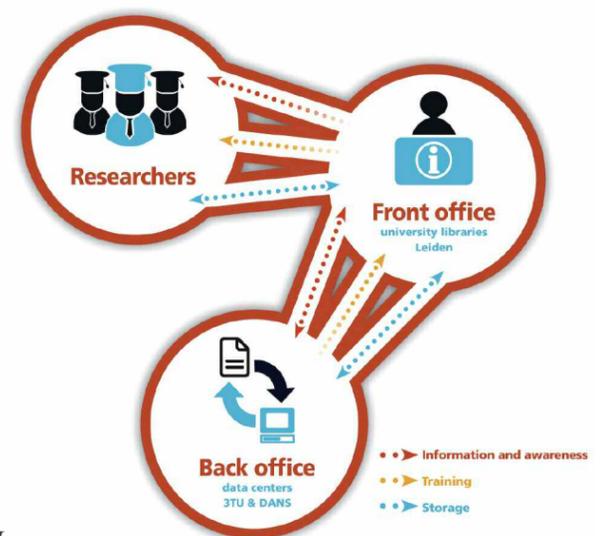
credit RDNL

Sharing data: good for science, good for you >>

6. So durable and accessible data storage ought to be the norm really. It's a way to add longevity and value to your data. Still some scientists have their doubts - but why? "A lot of researchers are concerned that others will make off with their data. Take me, I have stored 4,000 catalogued photographs here, with a total of 20,000 parameter values. You might use it to write an interesting article, so why don't I write it myself?"

7. Marion Wittenberg, data manager at DANS: "After storing your data at DANS, you can decide how others may access your data."