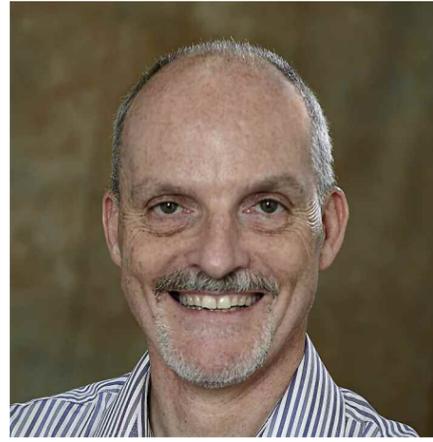




Christine L. Borgman photo UCLA



Katy Börner photo edu.com



Andrew Treloar photo Treloar



Herbert van de Sompel photo Elena Giglia

What does the world look like when sharing data is common practice?

# The vision of visiting fellows

Every year DANS invites visiting fellows to contribute to the Dutch research data landscape by sharing their knowledge. In this article they answer three questions:

1. Looking at your own daily scholarly practice, what is the aspect which frustrates your work the most, which you strongly dislike?
2. And what do you find most exciting?
3. If you were completely independent and able to create your ideal science/research lab – what would it look like?

## Christine L. Borgman:

“The ideal lab is one that brings the many stakeholders together”

1. I most dislike the administrative work, especially the lack of staff in public universities to handle some of the most basic duties. Then I remind myself that I am privileged to hold a professorial post in one of the world’s finest public educational institutions, the University of California.
2. Most exciting is the opportunity to influence practice and policy in the management of research data. My research and teaching interests have converged in the study of data and data practices, exploring how observations, models, artefacts and software become data; how these practices vary by individual and by discipline; and how these findings can be employed in the design of data collection, data management, data archiving and science policy. Factors contributing to the value of research data include the transition from print to electronic publishing, the ability to acquire and analyse large volumes of digital content in the sciences and humanities alike, and policies that promote openness and transparency.
3. On the surface, open access to data appears to offer vast benefits for research, education and innovation by leveraging public investments in these areas. Public policy documents suggest that releasing data

is an easy task to be accomplished at the time of publishing articles or books, and that research data are yet another genre to be absorbed by libraries and archives. Underlying these simple claims is a morass of theoretical, social, policy and practical problems. This morass has proven to be fertile ground for research. The ideal lab is one that brings the many stakeholders together to explore, discuss and experiment with new models of data practices, management and technology. DANS may be that ideal lab.

@SciTechProf

## Katy Börner:

“Let’s not waste expertise and time”

1. Many research teams that conduct science, technology and innovation studies use either Elsevier Scopus data or Thomson Reuters Web of Science data. The datasets are extremely comprehensive – covering much of humanities scholarly knowledge – but they are not perfectly clean. That is, all research teams apply their very own procedures to identify unique author names and institutions, to resolve errors in journal name and address data, etc. However, none of the teams is allowed to share cleaned data. That is, much of expertise and time is wasted cleaning data again and again. Even worse, teams cannot replicate each other’s results as they cannot access each other’s data.
2. More and more datasets are becoming available as part of the Linked Open Data effort ([lod-cloud.net](http://lod-cloud.net)). Institutions are making their own high quality data on faculty and their publications, funded research projects and courses taught available via research networking systems such as VIVO, Profiles NRN, SciVal Expert or CONVARIS. See the international network of emerging NRN systems at <http://nrn.cns.iu.edu>. Many of

these systems are compliant with VIVO ontology, i.e. they can be cross-searched to find research collaborators, mentors and reviewers.

3. R&D today is interdisciplinary and global. Novel tools and team management approaches are needed to assemble winning teams dynamically, to make them work productively and to diffuse results widely. Systems and research results presented at the joint [vivoconference.org](http://vivoconference.org) are highly relevant for making this work.

@katycns

## Andrew Treloar:

“I want to contribute to a better world”

1. The thing that I find most frustrating about my daily practice is all the things that prevent me from getting into, and staying in, a state of productive ‘flow’. I understand that constant interruptions from staff, phone calls and emails are part of the job, but on some days I cope better with this than on others.

2. The challenges that face the world are complex and scary, and what is needed to solve them is better knowledge and political will. I can’t help with the latter, but I can

help with the former. Meeting the challenges of managing and

working with data in support of our researchers is my way of contributing to a better world. And the fact that I find this intellectually stimulating is just a bonus!

3. Ah – an easy question. Cough. The ideal research environment for me would be one where the barriers to collaboration are reduced as far as possible, where the ability to build on previous work by anyone is maximised, and where researchers are able to spend their time focussed on making new discoveries rather than endlessly applying for research grants from a shrinking pool, or

publishing just for the sake of building their CV. Describing this ideal is of course very different to making it happen. Because of the lack of political will alluded to earlier, I fear that we need the impact of humans on the planet to become undeniable before society will commit to resourcing researchers more appropriately. Let’s hope that by that time it isn’t too late.

@atreloar

## Herbert van de Sompel:

“I am excited when I see efforts that fully embrace the web”

1. I work at an institution that does open and classified research. While the policy constraints that rule classified work are very stringent for obvious reasons, unfortunately, also open research is subject to strict policies that sometimes make things that are taken for granted at other research institutions difficult / impossible, like certain cloud services.

2. I am excited when I see efforts related to research and research communication that fully embrace the web rather than devise a research-specific enclave that merely uses the web as a conduit. Recent examples include the W3C PROV work and the Open Annotation effort that recently became an official W3C work item. I think the ongoing Research Object work is moving in, what I would call, the right direction too.

3. I am extremely intrigued by the notion of using global sync&share technology such as DropBox as fundamental research infrastructure. Currently, those technologies are missing features that are crucial from a research communication perspective. But, researchers absolutely love DropBox for its sharing, mobility, and versioning features. Architecturally, sync&share technologies blur the boundary between local file systems and the web; every file potentially has a URI. The result is a technology that allows every asset to seamlessly move into the web flow when so desired. In that regard, in my opinion, DropBox scores better than some special-purpose infrastructures. Hence, to me a crucial question is: How could we overlay the features that are missing from a research communication perspective on sync&share technology?

@hvdsomp

## FUTURE

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



13. Martijn Kleppe: “I estimate that 80% of all data are collecting dust in drawers or are dying on a hard disk. Along dies the information it contains, so a lot remains to be done.”

14. Sharing data: it’s good for science and for you. Please watch this movie at [www.youtube.com/user/DANSDataArchiving](http://www.youtube.com/user/DANSDataArchiving)