

# Mapping ownership in the data landscape

Seminar report:  
Research Data Ownership  
Rotterdam, December 2, 2014  
By: Aad van de Wijngaart

It's the oldest legal conflict in the world: who's entitled to knowledge? When Adam and Eve tasted the forbidden fruit of the tree of knowledge, God expelled them from Paradise and that's where all our problems started.

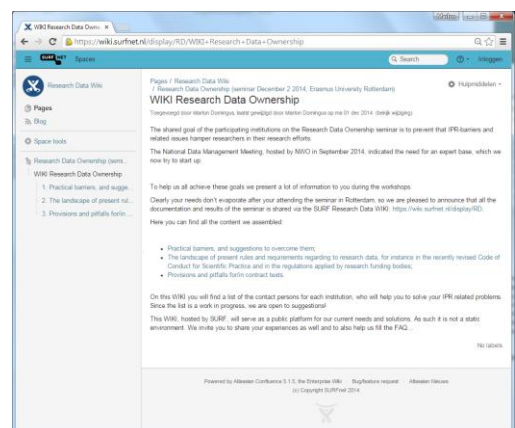
In science, the issue has become rather more complex. Many parties are involved in the creation of research data: not only the researchers themselves, but also the universities that employ them, plus funding agencies or commercial partners, and indeed the governments or intergovernmental bodies that make research policy. Trends in funding and the development of a powerful ICT infrastructure has helped to increase the sizes of both datasets and research groups, by involving multiple institutions, often in multiple countries. Naturally, all of these parties want to have a say over the fruits of the research.

Increasingly, Open Access is demanded for all of humanity: not only access to publications but also to the underlying research data. This would boost the pace of science and guarantee the most effective use of scarce research funds. These demands for openness and transparency have received a powerful boost from scandals about data that were falsified or even made up.

All this leaves researchers in a quandary. 'Publish or perish' is the mantra of the academic world. How can scientists be successful in this race when the data that they've worked long and hard to collect may be freely accessed and used by their competitors? How can scientific integrity be controlled without hampering the flow of research? And, what to do when openness conflicts with confidentiality, or valorisation with privacy? Is there really a way for researchers to protect their intellectual property rights (IPR) while sharing their data with others?

## Purpose of the seminar

The Research Data Ownership seminar was organised to help solve the quandary of how to protect your research interests when many parties want access to the research data. And, just as importantly, to show how you can get help from experts.



The organising committee consisted of a dozen people from five universities plus VSNU, SURF and DANS. They made use of a public WIKI created and hosted by SURF, dedicated to this very issue for the Dutch academic community. By doing so they contributed in accordance with one of the main conclusions of the National Data Management Meeting, hosted by NWO only three months earlier, namely the need for a central national expert base.

But the WIKI can only be successful as a shared responsibility of the Dutch academic community. Therefore the seminar was organised to inform, raise awareness and collect input for the WIKI from people in the field.

### **Program**

On the morning of December 2, 2014, a hundred participants from thirteen universities and numerous other organisations assembled at the Erasmus Centre for Entrepreneurship in the Rotterdam Science Tower. Nearly all of them were policy or support staff, including a number of legal advisors.

To kick off the seminar, first the rector of Erasmus University Rotterdam, prof. dr. Huib Pols, and then the keynote speaker, econometrist prof. dr. Dennis Fok presented several challenges that researchers and institutes face in their management of data. Pols emphasised the need for cooperation; Fok for professionalism. Both deemed it essential that clear and established practices are developed, with support from specialists in legal issues and data management.

During the lunch, participants had ten minute 'speed dates' with several such specialists, to discuss issues of privacy, Open Data, data management plans, and intellectual property rights.

The seminar concluded with an expert panel discussion, but that was only the dessert of the event. The main dish was the three workshops. Each of them covered one of the aspects that are to be addressed in the WIKI:

1. Practical barriers, and suggestions to overcome them (drs. Marlon Domingus, Erasmus University Rotterdam)
2. Relevant content on the rules and requirements of the main stakeholders providing research funding and grants (Esther Hoorn LLM, University of Groningen)
3. Provisions and pitfalls for/in contract texts (Rob Posthumus LLM, Erasmus University Rotterdam).



A wide range of issues, regulations and possible solutions was brought forward, by the three workshop leaders and participants. In the following paragraphs a number of these are presented in thematic order.

### **The 'default case'**

The most basic case that was looked at during the seminar is that of a PhD student who collects data for his thesis, without confidentiality or privacy issues, at a university that has no policy for data ownership.

Can the PhD do what he wants with the data, including taking it with him when he finds another employer after grading?

The discussion during the workshops showed that even this basic case is full of ambiguities. According to some, the PhD has the copyright, while the university holds the databank rights. The VSNU Code of Conduct also applies.

To complicate things, provisions in the collective labour agreement (CAO) may be indirectly applicable. But not every PhD is an employee at his university. If grant providers such as NWO are involved, they bring their own provisions, which include the right to impose further conditions. Practices may also differ between disciplines.

These complexities are not just theoretical hurdles. It has happened that a professor wrote an article that was based on data collected by a PhD, without mentioning whom he got it from. One of the participants reported that the uncertainty had often created problems at his institution. "We always come to the same conclusion: that next time we should organise this better..."

It's easier said than done. Groningen University has developed PhD regulations that deal with data-ownership, requiring that data be made available for further research. But they've been deliberately formulated as an 'open norm' so that the data management plan of each project can arrange this availability in accordance with current practice of the discipline involved.



Workshop facilitator Marlon Domingus observed:

"No one has a complete overview yet: whenever a problem comes up, people look at only one aspect. We are now exploring the whole landscape. But a year from now, these basic cases won't be so challenging anymore."

### **Crossing legal borders**

Another case that was discussed concerned the European Union: the EU sets higher standards with regards to the ethics of data storage than we are currently used to in the Netherlands. By implication a university should have a data protection controller and is asked to anticipate future policies on privacy and data. How do you deal with the situation where there's a gap between norms stated in the EU versus the Netherlands?

An example may be Open Access to data. What if the EU demands it, but the research consortium is led from the Netherlands and the data is collected in Africa by French researchers and stored in the United States?

One of the participants commented that this is the kind of problem that frightens researchers away from big programs, especially when there hasn't been previous cooperation with the

partners involved. Legal tools may help. Especially since researchers are not legal experts but want to stay safe.

### **Sharing with partners**

Another issue: five research groups work together collecting data, but only two of them are writing about it. There's a growing trend to view data as an independent research product. Is it possible to divide ownership of the data based on who collected what? Or should all partners be mentioned as co-authors of any articles that are based on the data?

Comments tended to emphasise the need for detailed agreements at the start of the project. Work-packages should be defined, that include who can publish about what. Essentially this isn't a legal but an organisational issue. If the arrangements are clear, there won't be legal complications, or so people thought.

Other commenters saw it as an ethical issue. Ethical norms can differ between national cultures, and who's to judge such an issue when trust is lost? It's difficult to deal with every possible conflict of interest when you start a project. But it may be possible to develop best practices and collect them in the WIKI.

### **Private partners**

Working together with private companies presents challenges of its own. In the third workshop, contract provisions drawn up by Erasmus MC were presented as an accepted practice of how to prevent conflicts of interest. These provisions are based on the European Horizon 2020 arrangements defined as an example of how to set up a collaboration. (The EUR provisions are also in the WIKI.)



As an example, it's important that negative results, too, may be published. Draw up a list of every acceptable reason for non-publication. Private partners should be allowed to comment on the draft of a publication, but only within thirty days. Some companies claim 120 days, but that's an unnecessary delay: big companies can move with great speed when they want, for instance when there's a patent to be claimed...

Universities on the other hand are not so fast. So be careful when a company wants you to bill all expenses within thirty days.

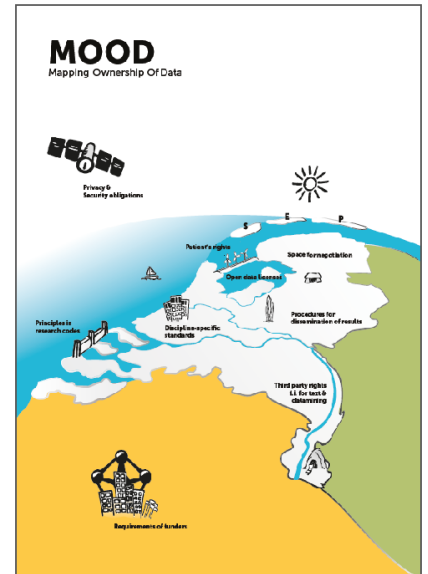
Some general advice from workshop facilitator Rob Posthumus: try to deal directly with decision-makers, preferably at the head-office of the company, if you're important enough as a partner. And do check the signature proceedings of your institution! Or you will learn *very* fast, but you won't be happy.

## MOOD Board

Decisions about ownership of data depend on the actual stage in the data life cycle. To make this important consideration as clear as possible, a 'MOOD Board' was thought out by Esther Hoorn and used during the seminar. MOOD stands for Mapping Ownership Of Data. The board is a visual tool, since scientists are humans too, and thus susceptible to visual stimuli.

On a map of the Netherlands, icons can be placed that represent relevant actors and factors, such as Open Data, Code of Conduct, requirements of funders, and discipline-specific standards. The data itself is represented by an icon of the sun, passing by these elements on its journey from dawn to dusk.

The MOOD Board is very handy for workshops and such. It can be downloaded from the WIKI pages and it's free for use under Creative Commons License.

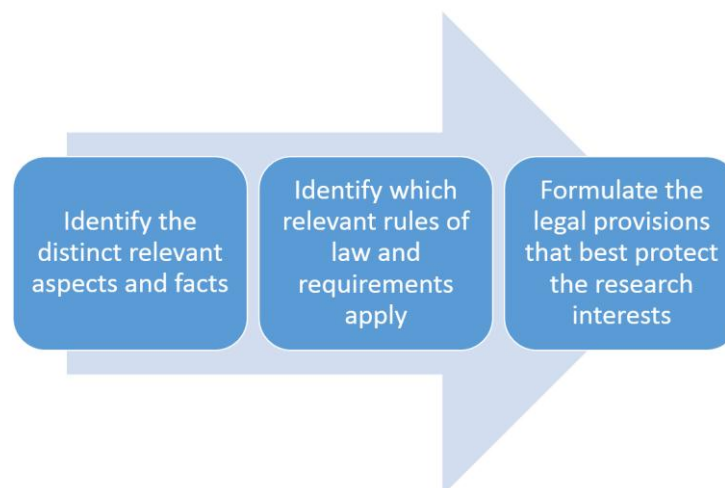


## Flowchart or matrix?

Another visual tool was discussed, showing principles, responsibilities and administrative/organisational elements that must be dealt with. It soon became clear that reality is quite complicated. For instance, flowcharts would need to be different for qualitative and quantitative research, and for national and international projects. They wouldn't only have to specify which regulations are relevant, but also how to deal with them.

The participants expressed a clear desire for a matrix (or matrices) where you could check what is and isn't applicable in a particular case. It should definitely include the stakeholders, such as the university and funders.

In essence the flowchart and matrix are decision support tools, supporting the generic model for identifying and qualifying IPR related barriers in a specific case and providing suitable legal provisions:



## Expert panel

After the main results of the workshops had been presented in a plenary session, it was time for the panel discussion.

The panel members were:

- Chair: dr. Matthijs van Otegem (Erasmus University Rotterdam)
- Prof. dr. Franciska de Jong (NWO board member)
- Drs. ing. Jacqueliijn Ringersma (Wageningen University)
- Esther Hoorn LLM, University of Groningen



Franciska de Jong shared her experiences as a researcher: she always got advice on data ownership when she asked support staff, but a lot is changing in this field and sometimes the support staff simply didn't know. The diversity in the scientific world definitely makes things more complicated: individual researchers often don't realise just how different other research communities can be. So any questions about intellectual property rights must first be placed in the proper context: there's no 'one size fits all'.

This is also why, she declared, NWO won't try to make everybody follow the same rules with the templates that it will design for data-management.

Jacqueliijn Ringersma agreed that right now, support staff like herself simply didn't have sufficient information to give advice with confidence: a lot is going on and there's a lot to find out. But the workshops had presented valuable information that she could use.

She also emphasised the importance of making good arrangements at the start of a project, since many problems are not legal but organizational. Scientists don't like to have lengthy discussions about these matters, but Esther Hoorn testified that research can be a lot more fun when uncertainties are out of the way.

The idea of creating a matrix of stakeholders and their relevant regulations was supported by Ringersma: it might be a good portal for the WIKI. At her university, staff from various units support scientists setting up projects. With the help of a matrix that support could be organised quite well. They were already using the WIKI and hoped to add to it.

Matthijs van Otegem observed that the audience probably represented a large part of the relevant support staff of the universities. Based on the positive reactions, he expected a follow-up seminar next year.



This Seminar report is funded by:  
DANS – Data Archiving and Networked Services  
PO Box 93067  
2509 AB The Hague  
T 00 31 (0)70 3494450  
info@dans.knaw.nl  
[www.dans.knaw.nl](http://www.dans.knaw.nl)