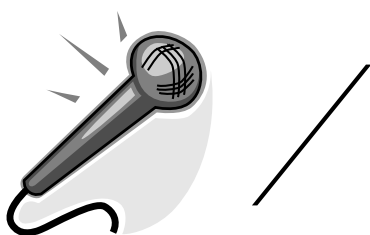
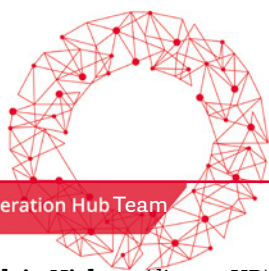


# A Conversation with the Team Behind the Data Federation Hub (DFH)



The Data Federation Hub (DFH) was launched last November under the auspices of Rector Elmer Sterken. On its website ([rug.nl/dfh](http://rug.nl/dfh)), the DFH presents itself as an organization connecting and building on several initiatives regarding research data at the University of Groningen (UG) and the University Medical Center Groningen (UMCG).



## Data Federation Hub Team

**Marjolein Nieboer** (director UB) \

**Ronald Stolk** (director CIT) |

**Teijo Doornkamp** (team member UB) /

**Anna Salzano** (team member CIT) \

**Salome Scholtens** (team member UMCG) |

**Marijke Verheij** (team member CIT) \

**Marina De Giorgi** (communication liaison CIT) /

**T**he DFH facilitates access to the expertise, support and facilities that researchers need to manage their data. It stimulates sharing lessons learned and propagates best practices throughout the organizations that support our researchers. It fosters the collaboration of data experts across disciplines and thus facilitate research.

What has been accomplished so far, and what are DFH's plans for the future? Pictogram spoke with the team that initiated the Data Federation Hub and learned about the past, present and future of the hub and the people involved in the initiative.

## Past

The DFH is an offshoot of the Human Subject Research (HSR) project. HSR is a two-year initiative in which the UG and the UMCG have joined forces to develop the tools, expertise and infrastructures needed to safely and securely manage sensitive human research data. The DFH wants to extend and consolidate the synergy among the partners in the HSR project, and to facilitate access to the data sources, tools and expertise on research data management that are currently scattered around the UG campus.

The DFH was initiated by the Center for Information Technology (CIT), the University Library (UB) and the Research-IT programme of the UMCG. But the driving forces behind the DFH are the groups that help researchers daily to improve the use and management of their research data.





*From left to right: Anna Salzano, Ronald Stolk, Salome Scholtens, Teijo Doornkamp en Marjolein Nieboer*

As Teijo Doornkamp explains: 'The Research Data Office (RDO) helps researchers from all faculties with their data needs. We have seen a lot of good practices and local solutions that are fit for use. But we've also seen bad practices and researchers struggling with their data needs. The RDO facilitates the exchange of solutions and best practices at the UG and the UMCG. The DFH strengthens that process.' Salome Scholtens affirms this: 'At several places at the UG, data managers or IT consultants already support researchers, but often in isolation. Expertise and solutions are not often shared.'

Marijke Verheij explains: 'The use of data, particularly "big data", is increasingly important in all aspects of science and society. In Groningen we manage and use huge, complex data collections for scientific discovery in several fields, for example radio astronomy, genetics, economics, computer science. Yet access to this data and to the expertise needed to manage it and extract useful information from it, is dotted around the campus.'

Researchers lack a comprehensive overview of what is available in terms of data, IT infrastructure, and training in this field. Anna Salzano: 'Initiatives concerning research data are sprouting up all over the Groningen campus, which is great. But there is no general awareness of what is available, and the links between these initiatives are often left to chance.' This is what the DFH aims to address.

### **DFH Meetups**

Since its inaugural meeting in November 2017, when Rector Elmer Sterken unveiled the identity and the logo of the hub, the DFH has organized a series of meetups to facilitate sharing lessons learned and propagate best practices. These events are widely advertised on the Groningen campus and are open to anybody who has an interest in the subject, or simply wants to learn more and connect with other research support staff.

During the meetups, research support staff can showcase what they do, voice their needs, network with colleagues and discuss how they can better support innovative research. How are the topics for these meetups chosen? Marina De Giorgi explains: 'We are open to any research data-related topic that is of interest to research support staff, and we welcome input from the audience. The line-up for 2018 is the result of input from participants at the DFH launch back in November 2017.'



*Marijke Verheij*

*Marina De Giorgi*

The February meetup focused on the storage and archiving solution for research data that is being developed for UG and UMCG researchers on the basis of the open-source platform iRODS. Jason Coposky, executive director of the iRODS consortium, was invited to speak, as well as colleagues from the University of Utrecht, where such a solution has been adopted.

The April meetup focused on training for research data management. Two experts came to talk about FAIR activities in DTL and ELIXIR-NL, and several UG and UMCG colleagues shared their experiences and initiated collaborations. Several support organizations committed themselves to jointly developing new training materials, doing more together than they could have done alone.

A fourth meetup in June focused on another 'hot' topic – Data Science. The keynote lecture was given by a representative of the Netherlands eScience Center. Experts from several disciplines – from linguistics to engineering, as well as experts on advanced methods of data analysis, such as machine learning and deep learning – shared their work and connected with colleagues.

The next meetup is scheduled for 18 September and the topic will be research data cataloguing. On 4 December a major event will be held to launch the tools and services for Human Subject / Sensitive data management that the HSR team has been developing for the past two years.

**Present**

The first meetups were attended by a diverse audience of about 50-80 people who welcomed the opportunity to connect with colleagues. Salome: 'Right from the start we wanted to make sure that everyone involved in research data support feels part of the DFH and is encouraged to give input. We want all our colleagues to see the Hub as an opportunity to showcase what their organization does, what data solutions they have, what expertise they can share, and learn from colleagues. We encourage them to share what tools they are building, and start collaborations to work on new ones together with other groups.'

Teijo explains that the DFH does not create things itself, but links together existing initiatives: 'DFH is a connecting platform, a hub consisting of the "spokes" involved. We do not aim to centralize or concentrate, but to link existing initiatives in research data management.'



‘Basically,’ Marijke adds, ‘we want to share what we already have and what we know. By improving how we manage our data we can achieve even more innovative science. There is no wrong door policy. All the data support organizations work together towards better research with better data.’

Organizing meetups on research data topics is just the start of the DFH activities. The DFH team plans to further develop the website as an online resource, a searchable catalogue where all data expertise, technical infrastructure for data handling, data catalogues and training courses on data stewardship and data science are made visible and accessible. Anna explains that a first version of a ‘portfolio’ tool is being developed, where the available tools and services for data management are listed by the service owners themselves and made available via a web interface. ‘We would like the DFH site to become a virtual meeting place for everyone involved in data management, a place where communities are formed and information is shared.’ Marijke: ‘The benefits of such exchanges can be far-reaching, and go far beyond the influence of the people who have initiated the DFH, resulting in new facilities and potentially revolutionary scientific discoveries.’

As Salome explains, the DFH is still in start-up mode: ‘We are using the resources we have to launch the platform and draw up an overview of what is needed, but we are aiming to broaden our base and take the DFH vision to the next level.’ At the moment there is a vacancy for a trailblazer (“kwartiermaker”) who will further boost the collaboration of the partners involved, operationalize the catalogues and the meetups, strengthen the brand and create a long-term plan. ‘It will be hard work, and the trailblazer will need the support of all our organizations. To be able to remain flexible and connect to researchers, however, we must keep the organization of the hub light and agile.’

Via the Data Federation Hub, researchers and research support staff and partners can gain access to the growing and extensive network of advanced expertise in all aspects of data management and data science. They can benefit from the advanced technical infrastructure, including virtual research environments, advanced software applications, optimal storage and archiving solutions, and access to high-quality HPC datacentres, capable of hosting vast amounts of data and computing power.

They can gain access to data sources by consulting the extensive FAIR data catalogues spanning several research disciplines, extending their reach nationally and internationally. They can gain access to the multiple options for training in Data stewardship and Data Science technologies through the existing services of the UG and the UMCG.

## Future


So what’s in store after 2018? The DFH team asserts that the hub is here to stay. Visibility of research data initiatives for the whole University is badly needed, and both the Board of the University and the organizations involved are committed to invest in the long-term viability of the DFH. The Groningen approach to research data management can and should be an asset to the whole University. Teijo: ‘That’s why we need to strengthen the internal network, learn from each other and adopt a shared brand, and become a unifying name in the field of research data management and data science for all research domains, both at the UG and at the UMCG.’

And what about the rest of the world? Marijke explains that the DFH is connected beyond Groningen via research communities such as DTL and Elixir. ‘We learn from them, but we also show them what we have. A representative from

the eScience Center attended the last meetup. We learned what they have available in terms of tooling, and also showed them how they can use of our tools and expertise.’

There is also ambition to open up to the Northern Netherlands, beyond the Groningen campus and towards private companies. Marjolein Nieboer: ‘This is, for the time being, just an ambition. It will be up to the trailblazer to further develop a plan. Our primary focus right now is to interconnect internally and support the research community at the UMCG and the UG. However, in the current climate we must not forget that universities also have a social obligation to create an ecosystem in which everybody can learn and make the best use of knowledge.’

What the DFH is doing sounds timely and logical. So why aren’t other universities doing this, or why wasn’t it here before? Ronald Stolk answers this question: ‘Groningen is ahead of the curve, but other universities are also moving in this direction. Together with SURF we collaborate in similar initiatives at other universities and organizations like Statistics Netherlands (CBS) and ASTRON. We don’t see this as a threat but as an opportunity to connect the DFH with other organizations, to learn from them and showcase what we do, how we work and what our assets are.’

As Marina explains, the DFH is open to new initiatives and collaborations in the field of research data: ‘We keep receiving requests not only to join the current initiatives, but also to facilitate new ones. It is great to see researchers and research support staff benefiting from the connections created via DFH, whether this is via direct email, conversations at the meetups, or simply by accessing the website.’ 



- Website Data Federation Hub: [www.rug.nl/dfh](http://www.rug.nl/dfh).