

Case Study: University of Groningen

How Groningen is drawing on the insights gleaned from Data Monitor to shape strategic thinking



Research Intelligence

Data Monitor

Track and analyse your institutional research data



The datasets that Data Monitor uncovered have been imported into Pure and are now powering key decisions from researcher assessment to future data policy

In 2019, Elsevier launched a 12-month pilot for Data Monitor, a new solution that provides institutions with visibility of their entire research data output. The University of Groningen in the Netherlands is one of just a handful of universities around the globe who joined the pilot as an early adopter. In this case study, two specialists from the library at Groningen reflect on their contribution to Data Monitor during its transition from pilot project to fully-fledged product. They also touch on the benefits they've gained and reveal how they plan to leverage their institution's datasets now they have been transferred to their research information management system, Pure.

Background

The past few years have seen rapid growth in the appetite for making research data publicly available. This has been led, in part, by mounting support for open science, concerns over research integrity, and the launch of initiatives such as the FAIR principles; a set of guidelines that seek to make “digital assets” Findable, Accessible, Interoperable and Reusable. At the same time, many want to see researcher evaluation expand beyond article and citation counts to include factors such as societal impact and their involvement in open science – including open research data.

In response, the number of open data repositories has risen sharply, along with open data requirements attached to research funding. And many involved in the research ecosystem, from policy makers and funders to publishers and institutions, have adopted new research data guidelines and practices.

The challenge

While many welcome this greater transparency, for those tasked with managing and analyzing their institution's research data output, the shift brings a unique set of challenges.

For example, most researchers deposit their data outside their institutional data repository. In fact, research* suggests that around 90 percent are hosted on one of the many external general subject or domain-specific repositories.

In the case of University of Groningen, with its rich history of empirical research, interest in research data really took root in 2013, as departments such as Genetics, Ecology and Economics began to increase their data production. Scientific Information



©UG, photo: Christina Elsenga

Specialist, Christina Elsenga, whose many responsibilities include providing research data management support, adds: “We had also experienced some research integrity issues. And researchers didn't necessarily feel we had the right tools to manage big datasets, so it was clear we needed to develop new ways of working.”

In 2015, a university-wide data policy was launched specifying that research data should be registered and made publicly available, whenever possible. And, in 2016, the University Library and IT experts collaborated on the launch of a research data office, now called the Digital Competence Center.

Christina says: “By 2017, my colleagues and I were manually finding datasets stored on external repositories and copying

* Zudilova-Seinstra, Elena; Zigoni, Alberto; Haak, Wouter (2020), “Analysis of research data for 11 Institutions - Data Monitor”, Mendeley Data, V3, doi: 10.17632/k5p45z33kb.3



the metadata to Pure. It wasn't an easy process. Researchers collaborate with colleagues all over the world so we knew the data could be stored anywhere. It took a lot of time to find queries/instruments we could use to help us scrape that information. But even then, we had to correct the metadata and import it. Often affiliation information wasn't available on the platforms we located and, if it was, it wasn't easy to search. Another issue was trying to ensure we reflected the wide range of research done here at the university."

She recalls: "It was the kind of work we had to fit in around the edges of working hours; for example, I should be cycling home but it's raining, I'll add a few datasets."

By 2019, Christina and team had added just over 600 datasets to Pure. She admits: "With all the copy and pasting, checking and improving on the metadata, it probably took us around 20-30 minutes a dataset – that's around 300 hours. We knew we had to find an automated solution."

©UG, photo: Egbert de Boer



“

Finding research data manually takes time – it's not sustainable. Data Monitor has automated the process, which is a huge step forward for us. And the link between Data Monitor and Pure is key.

— Christina Elsenga,
Scientific Information Specialist,
University of Groningen Library

”

Data Monitor





“The information that Data Monitor provides will really enrich the quality of the work we can do and support us in looking forward.”

— Cristina Montagner,
Data Consultant,
University of Groningen Library

At the same time, University of Groningen Library Data Consultant, Cristina Montagner, was involved in Research Output 2020 – a two-year project designed to optimize the university’s use of Pure.

She explains: “We felt that our research outputs – our publications – had reached a very mature stage in Pure. However, the same could not be said of our other data; for example, activities, research datasets and press/media coverage.

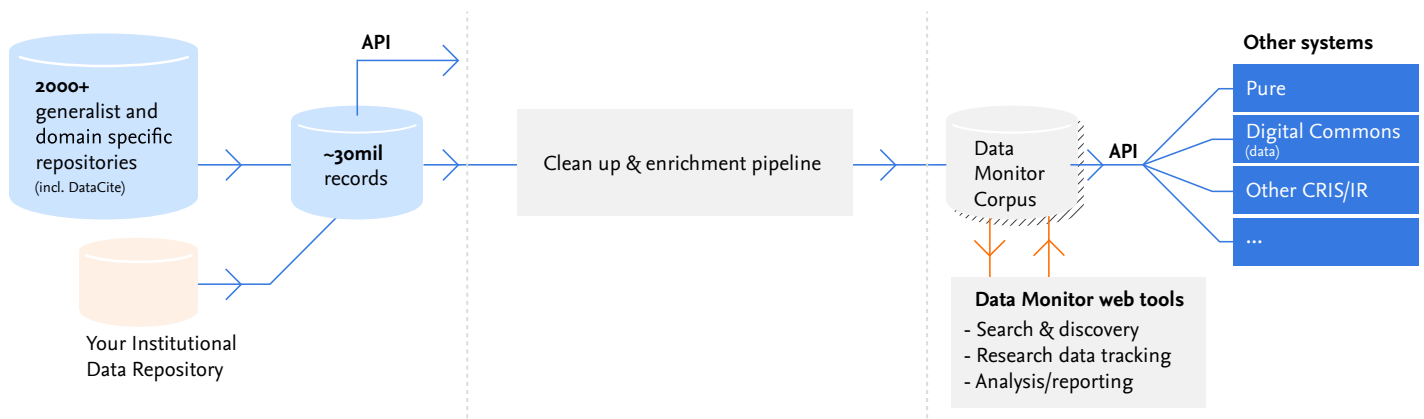
“And we know that research output should never be the only element of evaluation for a researcher, so we wanted to grow the range of information we could draw on and analyze.” She adds: “We also knew that there was a threshold we would need to cross – if we were going to say, ‘here are our datasets’ we had to make sure that what we were sharing was high quality and representative of all our disciplines and 11 faculties.”

The solution

While attending a Pure International Conference 2019 in Prague, the Czech Republic, Groningen’s Pure Application Manager, William Bargmann, heard about the Data Monitor pilot and plans to integrate the final product with Pure via an API. Cristina recalls: “We soon realized the pilot would be a great fit with our Research Output 2020 project and so we decided to get involved.”

While Cristina and William concentrated on the technical aspects of the Data Monitor pilot, Christina came on board to look at the functionality from a researcher perspective and validate the data quality. Any potential areas for improvement she identified, were fed back to the Data Monitor team.

How Data Monitor Works



Data Monitor: Taking the hard work out of research data management

By drawing on ~ 30 million research data records from more than 2,000 generalist and domain-specific repositories, Data Monitor can find, clean, enrich and deliver your institution's entire research data output. That information can then be used for analysis, reporting and showcasing, as well as making informed decisions in the RDM space.

Step 1:

Data Monitor locates and collects datasets that your researchers have published in open data repositories, both inside and outside your institution. These datasets (incl. source metadata) are accessible via a free API.

Step 2:

Data Monitor performs a clean-up, i.e., normalizes the metadata and removes non-reliable/non-data sources, dataset duplicates, datasets with dead links, non-data and SPAM records.

Step 3:

Data Monitor enriches the metadata of datasets with author/institutional affiliations (such as Scopus and SciVal Author and Institutional IDs) and links to publications.

Step 4:

The cleaned and enriched metadata is available in the Data Monitor Corpus, which can be accessed in two ways:

- Via the Data Monitor web application. With this option, it's easy to track your institutional datasets and monitor compliance with institutional RDM policies and funder mandates. Importantly, you can also expand your analysis beyond the boundaries of your institution.
- By integrating a Data Monitor API with a solution of your choice; for example, a research information management system like Pure. This option not only enables you to analyze, report on and showcase your research data, it also

makes it easy for others to find and reuse. In addition, you can maintain standard approval processes.

Data Monitor is integrated with Elsevier's Research Intelligence portfolio.

- Integration with Pure means you can automatically import dataset metadata and link it to specific people, organizational units, publications and other research entities. You can also showcase it in your Pure Portal.
- Integration with Digital Commons (and Digital Commons Data) ensures that the datasets your researchers publish in domain repositories are automatically featured on your institutional showcase page.
- Integration with SciVal enables research administrators to assess the impact of RDM best practices; for example, you can compare the impact of publications with datasets versus the entire institutional research output.
- Integration with Scopus means your researchers can find datasets matching a query that you used in Scopus when searching for research publications and other content.

In fact, for Christina and Cristina, the collaboration aspect of the pilot process proved particularly enjoyable, as Cristina explains: "I really appreciated the breadth of knowledge on the Elsevier team. There were people who could explain the technical reason why something was created that way, but also people who could explain what they hoped to achieve with it. That was important for us as users."

She adds: "We were quite active with our feedback. Our comments were accepted and, to be honest, the Elsevier response was often very quick. A good example is the synchronization functionality between Data Monitor and Pure. Elsevier offered us an option that would automatically update the data in Pure, but everyone on the pilot recognized that this wouldn't benefit our daily work. The Data Monitor team were surprised, but they listened, and we got an option better suited to our needs."



©UG, photo: Marcel Spanjer

Alongside University of Groningen, there were four other universities involved in the pilot: Vrije Universiteit Amsterdam, University of Southern Denmark and Belgium's Vrije Universiteit Brussel and University of Namur. For Christina, the opportunity to work alongside those universities was a highlight. "They've helped us gain new ideas and insights."

According to Lorenzo Feri, Director of Product Management at Elsevier, who was closely involved in the pilot, the feedback received from University of Groningen and the other participants proved invaluable. "We were really grateful for their honesty and many improvements they suggested have now been addressed. Even though Data Monitor has now been released as a product, we don't plan to stop development. On the contrary, we have a lot of data quality and integration improvements in the pipeline."

"It's not just about increasing the number of datasets in Pure by 10 or 1,000 – it's about the fact that they are now discoverable. We have a much fuller view of what's going on in research and so do visitors to our Pure Portal."

— Christina Elsenga

Results

The University of Groningen team has only just moved from pilot phase to a fully-functioning Data Monitor integration in Pure, and there is still some validation work to be done. But the number of datasets captured looks set to grow from just over 600 to 3,000. Importantly, the datasets reflect the university's full range of faculties and disciplines.

According to Cristina, now the step of collecting the data has been mostly removed, there is more time available to focus on leveraging the data to support decision-making. This includes:

- **Understanding faculty use of external repositories.** Cristina explains: "We can see if they are using them frequently or not. And if they aren't, why not? Maybe they need more support, maybe the kind of data they have are not suitable, or maybe they are still storing them on a USB stick! We can also check they are complying with legal requirements like GDPR [the EU's General Data Protection Regulation]."
- **Supporting the work of the Research Impact Services:** Cristina says: "There are opportunities to develop new analyses at the university, faculty, research group and individual researcher level."
- **Incorporating the data into the annual researcher evaluation process.** It will be used alongside alternative metrics, such as media mentions, to understand a researcher's societal impact.
- **Informing the university's future research data policies.**
- **Improving the data lifecycle support for researchers.** According to Christina, Data Monitor has revealed a whole raft of repositories that she was unaware researchers were using and plans are underway to use those insights to improve researcher support: "For example, we can make them more aware of their options, advise them on improving their metadata and develop a list of recommended repositories. We may also discover needs an external repository can't meet." Groningen is also working with the Pure and Data Monitor teams to refine a new feature it requested - the introduction of a common standard for dataset metadata which complies with Europe's OpenAIRE guidelines. According to Christina, she and her colleagues have been invited to test the integration and help iron out any teething issues.
- **Complying with the Netherlands' new Strategy Evaluation Protocol.** This policy was updated in 2021 to include open science and recognition and rewards for academics. Christina explains: "Research institutes and research groups will be officially assessed once every six years and carry out mid-term evaluations themselves. One of our faculties was evaluated a year or so ago and showcased some early work on datasets, claiming there would be more in the future. They were explicitly praised for that and now we can actually deliver on that promise."

These are only some of the early plans. Both are confident that new opportunities will emerge as they start to work more closely with the data. Cristina adds: “You can only take steps like these when you have data that are representative of your university’s situation. With Data Monitor we finally have that.”

About University of Groningen

University of Groningen is a Dutch research university with a global outlook and a rich academic tradition. Since it was established in 1614, it’s been recognized for its trailblazing approach; for example, it was home to the first female student, the first Dutch astronaut and it has produced numerous Nobel prize winners. It currently sits within or around the top 100 on several influential university ranking lists.

Geographically, the university is rooted in the Northern part of the Netherlands, a region very close to its heart. Its goal is to connect education and research with sustainable and economic processes within society. This comes together in its focus on Energy, Healthy Ageing and Sustainable Society research.

With its annual budget topping 900 million Euros, the University of Groningen is home to 11 faculties, 36,024 students enrolled on Master and Bachelor programs, and 3,400 full-time academic staff. Each year, it produces around 9,000 research publications.

“Now that Data Monitor is integrated with Pure, our Pure Portal is starting to show a much fuller picture of our research. It makes Pure more complete in all the aspects important to a modern researcher. And it’s visible for everyone to see.”

— Christina Elsenga



Data Monitor

Data Monitor

Data Monitor puts institutions in the driving seat of their Research Data Management journey by providing visibility on their institution's entire research data output.

Pure

The world's leading research information management system that enables evidence-based decisions, simplifies research administration and optimizes impact, reporting, and compliance.

For more information about Data Monitor,
visit elsevier.com/solutions/data-monitor

Elsevier offices

ASIA AND AUSTRALIA

Tel: + 65 6349 0222

JAPAN

Tel: + 81 3 5561 5034

KOREA AND TAIWAN

Tel: +82 2 6714 3000

EUROPE, MIDDLE EAST AND AFRICA

Tel: +31 20 485 3767

NORTH AMERICA, CENTRAL AMERICA AND CANADA

Tel: +1 888 615 4500

SOUTH AMERICA

Tel: +55 21 3970 9300



Data Monitor

Copyright © 2021 Elsevier B.V.
July 2021