

The Real World of Research Data



Connie Clare, October 2019

Colophon

Author/Illustrator

Connie Clare, TU Delft, The Netherlands, [0000-0002-4369-196X](#)
Author contact: Twitter ([@ConnieEClare](#)), [LinkedIn](#))

Data Champion Contributors

Faculty of Applied Sciences

Anton Akhermov
Victor Koppejan
Raúl Ortiz Merino

Faculty of Architecture and the Built Environment

Balázs Dukai

Faculty of Civil Engineering and Geosciences

Siân Jones

Faculty of Industrial Design and Engineering

Natalia Romero Herrera
José Carlos Urra Llanusa

Faculty of Mechanical, Maritime and Materials Engineering

Pavlo Bazilinksyy
Joost de Winter

Faculty of Technology, Policy and Management

Anneke Zuiderwijk

University Corporate Office, ICT Innovation

Susan Branchett

Licence

[CC BY 4.0](#)

Reference

Clare Connie. (2019). The Real World of Research Data [Book].
Zenodo. <http://doi.org/10.5281/zenodo.3584373>

Table of contents

Author/Illustrator	1
Data Champion Contributors	1
Licence	1
Introduction	3
Reduce, Reuse, Recycle knowledge: How open hardware can help to build a more sustainable future	5
Keep calm and go paperless: Electronic lab notebooks can improve your research	9
The Changing Landscape of Open Geospatial Data	13
Carve your niche with 'The Carpentries'	17
Empowering people with data	21
Starting small and thinking big: A Quantum Tinkerer's quest to mentor Open Science	26
Five W's of Open Data	30
Switch gear! Drive the uptake of Open Science within your research team	34
Helping researchers do IT	38
How to build a community of Data Champions: Six Steps to Success	42

Introduction



Researchers engage with research data in many different ways. We want to understand how researchers use and share data, and learn what good research data management (RDM) means to them. To explore the real world of research data, we hear first-hand accounts from our Data Champions who lead the way to Open Research at [Delft University of Technology \(TU Delft\)](#).

What is a ‘Data Champion’?

Inspired by the [University of Cambridge Data Champion](#) programme, TU Delft launched their [Data Champion programme](#) in September 2018 as a cost effective way to provide disciplinary-specific RDM support for all eight university faculties. There are currently 47 Data Champions volunteering their time, effort and expertise at TU Delft and the numbers are continuously increasing.



An individual who volunteers their discipline-specific expertise, promotes FAIR data principles and advocates proper research data management (RDM).

An individual who uses their passion for knowledge exchange and their desire to build a collaborative and researcher-led community to drive the uptake of good RDM within their faculties and departments.

Why is this book important?

Funding bodies, political organisations and research institutions worldwide emphasise the value and importance of good RDM as a prerequisite for Open Science, but the practice is slow to change. Researchers may be reluctant to change their current data management practice if they feel that making their research open will add to their already burdensome academic workload. They may also feel vulnerable to being ‘scooped’ on prospective projects if they share their hard-earned results with others, or they may be simply unaware of the benefits of Open Science. What’s more, researchers may not feel incentivised to change their practice since Open Science isn’t clearly recognised by the traditional academic reward system.

Such barriers to realising good RDM and Open Science are further exacerbated by the fact that discussions are typically limited to librarians, data professionals and a handful of researchers who are already passionate about good data stewardship. Therefore, we believe that engaging researchers, as the main data producers and reusers, is critical to engender a culture change towards good RDM across an entire research institution.

Indeed, this book demonstrates that a culture change begins with individual researchers. By presenting real-world, first-hand accounts from researchers of different career stages and diverse scientific

disciplines, we show that fostering an open, collaborative and multidisciplinary research environment offers benefit for all. From using electronic lab notebooks to study the rheology of foam, to reusing geospatial data to build 3D city models, each article demonstrates that data intersects with research in a multitude of ways and that common foundations of research integrity unites all researchers.

Our Data Champions take leaps forward to advance open science by adopting good RDM and FAIR data practice. We believe that expressing gratitude to individuals who ‘go the extra mile’ and supporting them in their mission to champion Open Science will not only sustain motivation and productivity amongst the existing cohort of Data Champions but will incentivise others to join our vibrant community.

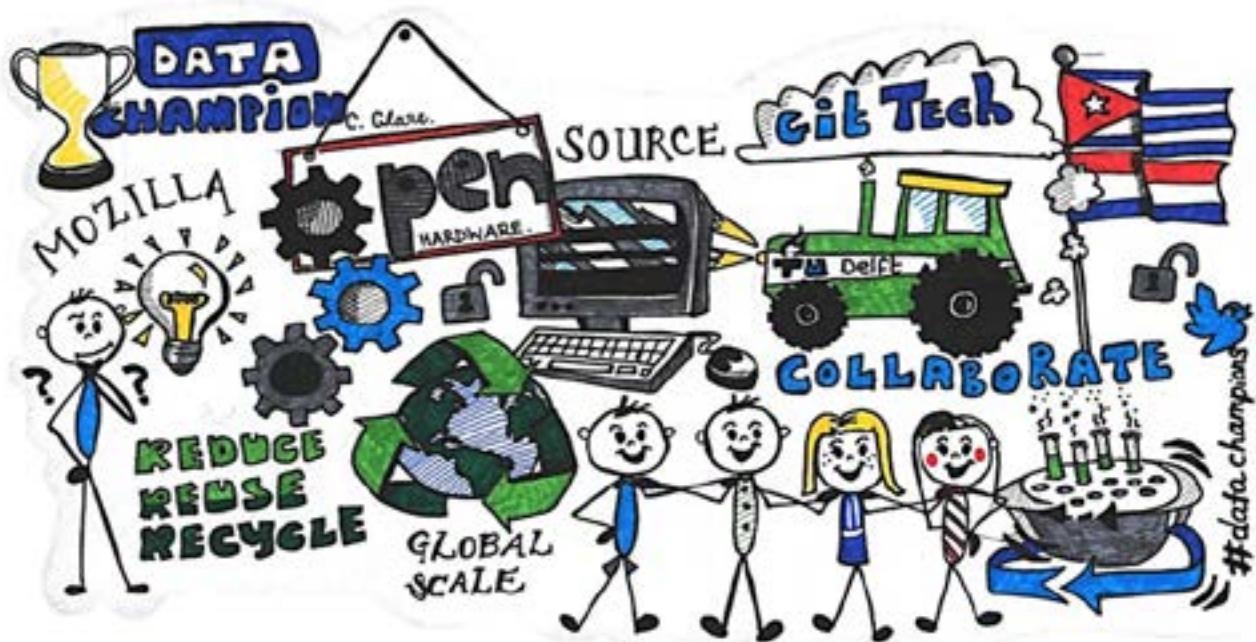
If you’re inspired by the work of our Data Champions and want to learn how to build your own community-based model within your institution, read our toolkit, ‘How to build a community of Data Champions: Six steps to success’ which can be found at the back of this book.

What is this book about?

This book presents a selection of articles to showcase the work TU Delft’s Data Champions. Our primary objective was to reward and recognise their exemplary efforts by sharing their success stories as written articles to increase their visibility and raise their professional profile. Each article has been written to inspire and inform the wider research community and to help drive a culture change towards Open Science.

From June to August 2019, one-to-one interviews were conducted with 11 Data Champions to learn: about their research; their motivations for becoming Data Champions; how they effectively engage with researchers; and, their future goals and aspirations. Following each interview, an article was written and published on TU Delft’s [Open Working blog](#) under a dedicated tab titled '[Data Champions](#)'.

Reduce, Reuse, Recycle knowledge: How open hardware can help to build a more sustainable future.





[José Carlos Urra Llanusa](#) from the [Faculty of Industrial](#)

[Design and Engineering](#) shares his vision of open hardware with TU Delft Library.

Industrial and interaction designer from Cuba, [José Carlos Urra Llanusa](#), recently joined our team of Data Champions at TU Delft to share his engineering and programming expertise. Here, he tells us more about his passion for using open hardware to promote sustainable development and to facilitate a circular economy where resources are not discarded but are repaired, improved and reused to reduce waste.

The benefits of open hardware

Open hardware means that the information about a product's design (e.g. blueprints, mechanical drawings, schematics and source code) is available so that others can study, replicate and improve the design. Since users are free to modify the source files and underlying code, there is an increased capacity to improve the functionality and physicality of a device, which in turn promises greater social impact. However, like software, hardware designs are subject to copyright and patent law that protects a design's core technology and prevents its reuse. "Hardware design is very rooted in a proprietary centered culture," says Urra Llanusa. "There's a conflict of interest and no simple solution but the fact is that we are more productive when we share and build upon each other's work. This is the point of science; to deliver replicable experiments and research that someone else can build upon."

To build an open future, we must understand the past

"Adopting an intellectual property-oriented mindset is not something I see as negative or bad; it's just a culture based on history," he adds. The message is clear that if we are to facilitate a transition towards an open culture, then we must appreciate the concerns of those who are unwilling to change old habits and cooperate with those who are prepared to tune into a new way of thinking. Urra Llanusa believes that "The success of open source inventions and technological advancements are obvious. We can learn a lot from the history of the internet since the software industry has been a driving force behind the open source developments."

Working with the future generation

With a strong interest in education, Urra Llanusa, is currently involved in the planning of a Circular Design course with aeronautical and industrial design engineer, [Dr. Ir. Bas Flipsen](#). Urra Llanusa also

runs workshops for MSc Engineering students at TU Delft to teach them the basics on how to use [Git](#) (a distributed version-control system for tracking changes in source code) and [OpenSCAD](#) (a software for creating 3D computer aided design (CAD) objects). Around 20 students from three Masters programmes ([Design for Interaction](#), [Integrated Product Design](#) and [Strategic Product Design](#)) attend workshops to learn how to work in collaboration using [OpenSCAD](#), [Git](#) and [GitHub](#). His workshop on [OpenSCAD is publicly available](#) and meant to be replicated, reused, improved by other instructors.

Urra Llanusa takes great pleasure in conducting student workshops. "It's so fun!" He remarks, "I've had so much positive feedback from the students. It would be great to implement these workshops into their curriculum in a more systematic way."

By documenting anonymised [feedback](#) from the students undertaking his workshops, Urra Llanusa has been able to gain valuable insight into their learning preferences and personal opinions on the implementation of open source tools during their degree programme. This work presents a shining example of how bottom-up teaching initiatives give early career researchers a voice to direct their independent learning goals.



Caption: Urra Llanusa conducts collaborative workshops on OpenSCAD and Git technology for MSc Engineering students at TU Delft.

A ‘hands-on’ approach: problem-solving with open hardware

Aside volunteering his skills and expertise, Urra Llanusa is invested in the development of [various open hardware projects](#) that have social and sustainable impact through problem-solving. He is a co-founder of [Go!Commons](#) (& open source things), a project that documents the designs and developmental processes of open hardware, such as agricultural and manufacturing machines. Since these designs are made publically available on [GitLab repositories](#), anyone can use and modify them to create working solutions. Click [here](#) to see the work that's been achieved!



Caption: An overview of the Go!Commons project on open hardware.

As a participant in the [Mozilla Open Leaders](#) programme, Urra Llanusa began this work studying a [microwave](#) with interest in open source households. He then decided to collaborate with communities on existing projects, focusing on the design methodology and best practices for teaching git and gitlab on those projects. He explains the aims of his latest project, the replication of a [wet-lab centrifuge](#). “By deconstructing and rebuilding the instrument alongside my collaborators, together we can learn how the instrument works. Suddenly, industrial products are no longer a black box.”

In addition, Urra Llanusa has worked as a designer and developer on the [Open Source Ecology](#) project, an initiative that aims to manufacture open source industrial machinery at a fraction of the commercial cost through global collaboration. He shared his inspirational story of helping a farmer in Maine, USA, to document and use git based workflows for his own CNC Plasma Cutter to build tractor implements – An example of using open hardware for self-sufficiency and economic sustainability.

A sustainable future

Following two rounds of serving as a mentor for the Mozilla Open Leaders programme, Urra Llanusa plans to launch his own Open Leaders programme through the next phase called [Open Leaders X](#). The course teaches participants how to conduct their own community-run Open Leaders programme and offers an exciting opportunity for leaders to learn and improve open hardware practices.

Urra Llanusa concluded our conversation on using open hardware to build a sustainable future with some words of wisdom: “The best things can happen spontaneously. You can have a goal or an idea but then you have to knock on doors and talk to people in order to build a network. People can work together to create a community; it’s important to engage with all kinds of people. – That’s worked for me.”

Keep calm and go paperless: Electronic lab notebooks can improve your research.





Postdoctoral researcher from the [Faculty of Civil Engineering and Geosciences](#), [Siân Jones](#), champions the use of electronic lab notebooks (ELNs) at TU Delft.

Physicist and [Data Champion, Dr Siân Jones](#), has dedicated her research career to studying the [rheology of foam](#) at several European institutions. She currently works in the [Petroleum Engineering](#) section at TU Delft studying the industrial application of steam-foam floods at high temperatures and pressures for enhanced oil recovery. As an experienced bench scientist, Jones gives us her verdict on the use of electronic lab notebooks (ELNs) to promote good research data management and laboratory practice.

Problems with paper...

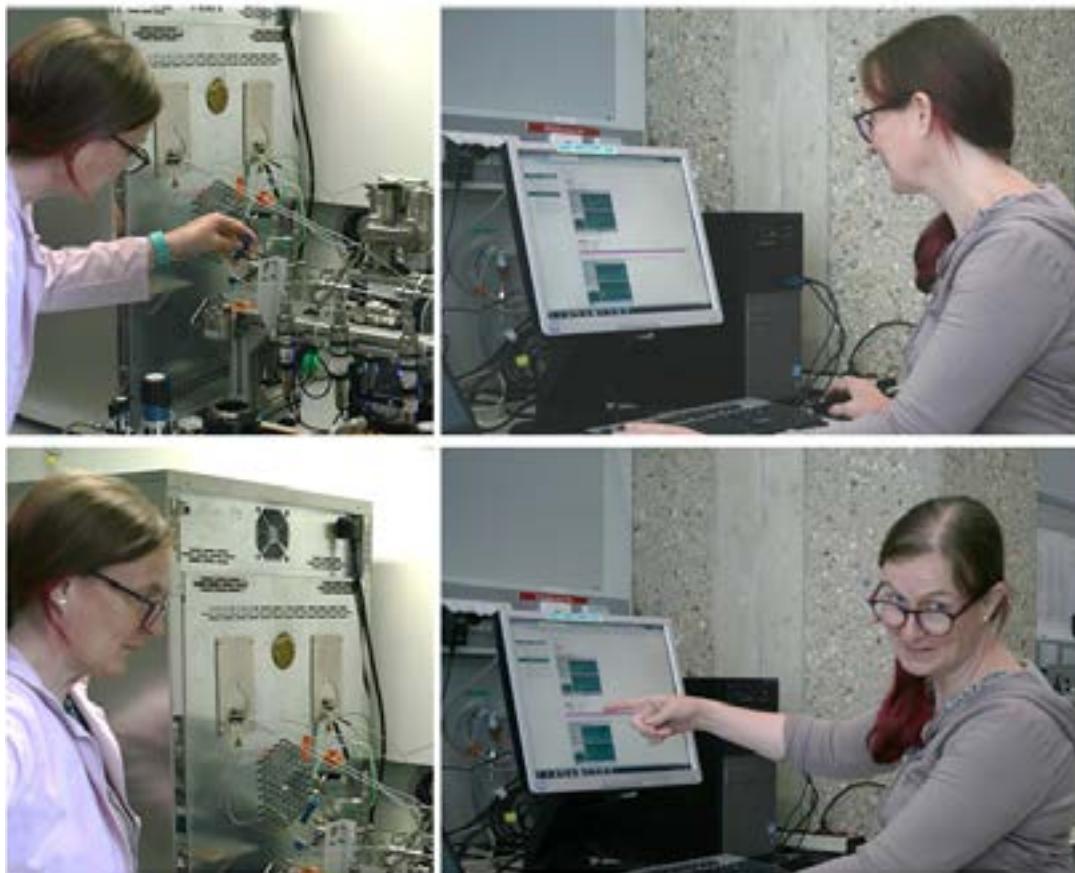
Life in academia can take a researcher from one short-term employment contract to another. As a researcher moves to the next institution, country or continent, it's important that they leave their data in a well-organised format that is secure and dependable for optimal knowledge exchange. "I've relocated for work on many occasions," says Jones. "Each time I've left an institution I've questioned the most effective way to archive my data to benefit my successors."

Jones recounts her frustration with relying on paper notebooks to record experiments. "It becomes problematic when you need to replicate experiments or continue a research project conducted by a researcher who has since left your institution without documenting their work properly." Jones identifies some difficulties of understanding handwritten information (including her own!) "Locating and interpreting data from past experiments can be challenging using paper notebooks," she admits. "Finally, after rifling through pages of scribbled diagrams, photocopies and post-it notes you find the experiment you've been searching for and it can be impossible to decipher the handwriting and bridge the gaps of missing information". These problems sound all too familiar to most lab-based scientists and have increased the demand for digital solutions, such as ELNs, that can improve the rigour, robustness and reproducibility of scientific research.

The electronic lab notebook

In its simplest form an ELN is a software system for documenting research work. Over time, they have evolved to become much more than a direct replacement for their paper ancestor, encompassing a range of additional features, including:

- A text editor for writing notes in a way that replicates a paper notebook
- Spreadsheet tools for calculations, and formatting tables and graphs
- Protocol templates for documenting standard procedures
- Laboratory inventories for documenting samples, reagents and apparatus
- Collaboration tools for sharing experimental information.



Caption: Jones demonstrates the use of an electronic lab notebook.

Based on Jones's experience of using various ELN software packages, including [BIOVIA](#), [labguru](#), [labfolder](#), [RSpace](#) and [eLABJOURNAL](#), we asked her to tell us why she advocates ditching the humble paper notebook to go digital. She explains how ELNs provide solutions to problems with paper by offering...

Structure (all of your data in one spot!)

Most ELNs have a user-friendly interface that allows the systematic input of data section-by-section.

Working with a digital template simplifies data curation and retains data in an organised structure in a single location. Plus, templates make it easier to create metadata for your experiments.

Bonus: Reusing digital templates for repetitive tasks, such as laboratory protocols, saves time and effort thereby maximising efficiency.

Searchability

Advanced search options mean users can retrieve information by 'author', 'tag', 'unique ID' and 'textual content' so that data can be located in seconds.

Bonus: You'll never need to worry about being able to read your own writing again!

Shareability

ELNs facilitate collaboration locally and globally by allowing users to share data. By encouraging open and transparent working, ELNs promote inclusivity as researchers can connect through the tool and extract the information they need. Some ELNs even have a live news feed to broadcast experimental results amongst laboratory group members!

Bonus: Technicians and PIs can offer their guidance, support and can track student progress. This is particularly useful when they are 'on-the-go' or away from the lab

Seamless integration

Most ELNs are accessed via a web-browser, which means they allow remote access from any location and are not restricted to one operating platform; several ELN providers have created apps to run their software on different device types (e.g. smartphones and tablets). Integration with web apps, such as [Microsoft Office](#), facilitates the quick and easy importation and exportation of research data. Moreover, ELNs are designed to connect with pre-existing [laboratory information management systems \(LIMS\)](#) for more effective data management.

Bonus: You can export your research data if you decide you want to disengage from a particular ELN software or if you want to switch to a new one.

Secure storage

Typically, ELN data is password-protected and backed-up to a cloud server for secure, long-term storage in compliance with good laboratory practice (GLP) guidelines, local data protection regulations (e.g. GDPR) and funding agency requirements. On-premises client-server installation is the most suitable option for institutions that must comply with strict IT policies to uphold data confidentiality. Electronic signatures allow for the validation of data integrity by authors and administrators. What's more, if all members of a research group use the same ELN software, it enforces a common language throughout the group which safeguards the body of information generated.

Bonus: Turn to a cloud-based solution to transform your disorganised stack of paper into a centralised data hub and you'll have more room in your filing cabinet and less lugging around of heavy lab books.

Implementing an ELN: Trials and training

The ELN market boasts considerable variety. [With more than 70 active products identified in 2017](#), the burgeoning array of choice can make it a challenging task to find the right ELN package for your research group. To avoid becoming overwhelmed by choice, Jones advises researchers to “only select two or three ELN packages to trial at any one time”. She highlights the difficulty of choosing a suitable ELN for deployment across all university departments, claiming that “there’s no one-size-fits-all solution. ELN designs are often tailored for use in specific research fields which can make choosing a suitable software highly subjective”. Most ELN vendors offer a free, cloud-based service for individual users, however, storage capacity may be limited and there’s minimal opportunity for collaboration with colleagues. Jones agrees that “ELNs are most valuable if all members of the research group use the same software”.

Once you’ve chosen your ELN, it’s important to provide adequate training for all users. “ELN software implementation takes time, commitment and adaptability,” says Jones. “It can be difficult to persuade established researchers to change old habits and start using ELNs but once it becomes common practice, the long-term benefits significantly outweigh the initial work investment.” As ELNs present a culture shock for some, for others it’s a welcomed change. Jones reveals that early career researchers, who have grown up in the digital age, tend to embrace electronic solutions and are eager to wave goodbye to their paper notebook.

Need more information?

For more information on the implementation of ELNs at TU Delft, you can read our blog post about the [library events that took place in March 2018](#) and our current [pilot project](#). Alternatively, you can talk to ELN enthusiast, [Dr. Siân Jones](#) or Data Steward and trials coordinator, [Esther Plomp](#). General information on the use of ELNs for improved research data management are provided by several other academic organisations, including [The Gurdon Institute](#) and [Harvard University](#).

The Changing Landscape of Open Geospatial Data.





"We can harness the power
of existing open data"
says [Data Champion, Balázs Dukai](#).

Software engineer and [Data Champion](#) from the [Faculty of Architecture and the Built Environment](#), [Balázs Dukai](#), specialises in the construction of geospatial datasets to generate 3D city models. A '3D city model' is a three-dimensional (3D) digital representation of an urban area that depicts its landscape and architectural infrastructure. The geometry and structure of the terrain, roads, rivers and buildings are described by 3D geospatial data (i.e. coordinates, projections and transformations) for their contextual visualisation using computer modelling software.

Data in all shapes and sizes

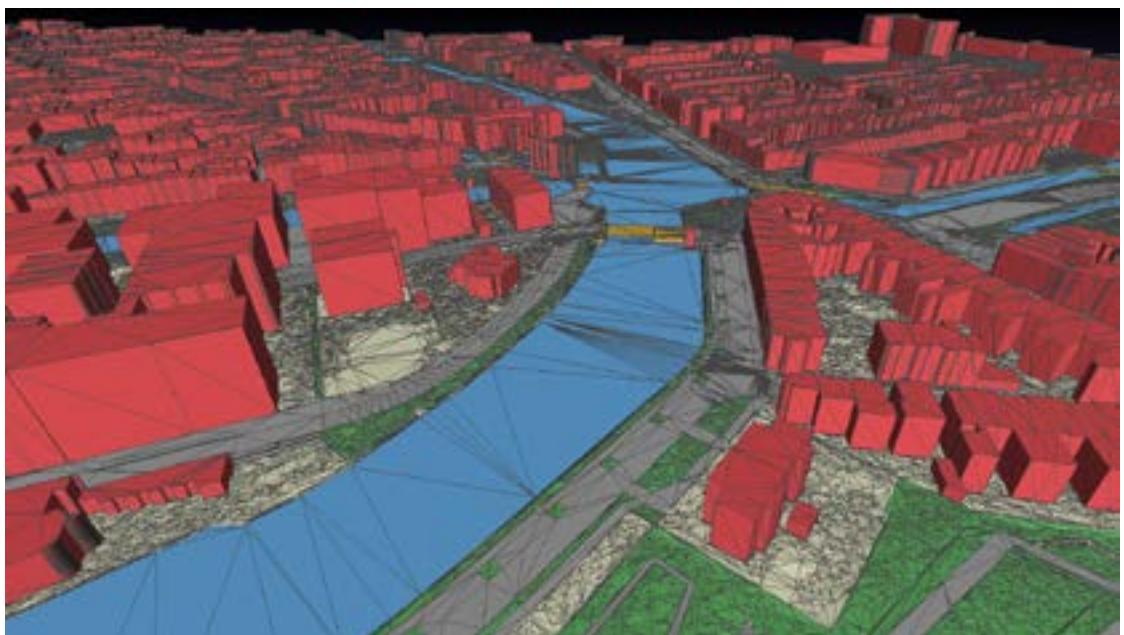
The ability to translate our physical world into a virtual reality has become a valuable asset in the planning, visualisation and management of a wide range of urban applications. By way of example, 3D city models have been used to estimate population size, energy usage, pollutant tracking, solar potential, shadow analysis, wind turbulence and noise impact. With so much to offer, it's no surprise that there are now [more than one thousand virtual city models](#) worldwide.

Rapid population growth and urbanisation is increasing the demand for more complex 3D models that can improve urban development. Dukai discusses some of the current issues surrounding the generation of complex geospatial data, "There are already numerous [open geospatial datasets available](#) for use but due to the heterogeneity and complexity in data formats, the exchange between them is difficult at best, but often not possible." He explains that this problem is confounded by the lack of software support for the utilisation of 3D geospatial data, "3D modelling software is limited to expensive proprietary platforms that benefit many commercial enterprises but few individual researchers." Dukai expresses his belief that "there is enormous value in existing datasets that are stored away in repositories collecting dust because they are not interoperable – it's such a waste."

Dusting off data

Harnessing the power of existing geospatial data to produce 3D city models has been at the forefront of Dukai's research agenda. At the [3D Geoinformation](#) research group at TU Delft, he is working on [3dfier](#), an open software that transforms 2D topographical datasets to 3D by lifting every polygon using height information from point clouds. The software combines a digital map of the Netherlands ([The Basic Registration of Large-Scale Topography, BGT](#)) with the digital elevation model for the Netherlands ([The Actueel Hoogtebestand Nederland, AHN](#)) to create a volumetric 3D representation of the urban environment. So far, 3dfier is used by the Dutch national mapping agency (Kadaster) to reconstruct the 3D model of the Netherlands — Check out the [animation](#) to discover how 3dfier was used to create Delft's 'digital twin'!

3dfier has also been used to create the [3D Registration of Buildings and Addresses \(3D BAG\)](#) using the [original Dutch BAG register](#) and the digital elevation model ([AHN](#)). This map provides details of a building's address, its construction and registration date, and its intended use. By clicking on a building footprint, you can learn more about a building's attributes, such as ground surface or roof height. The 3D BAG dataset is automatically updated each month to provide users with an up-to-date record of the latest building stock and elevation information as buildings are built, registered and demolished. 3D BAG data is [free and available to download](#) and can be accessed via our [4TU.Centre for Research Data repository](#).



Caption: An example of a 3D city model created by TU Delft software



Caption: 3D BAG is the most detailed, openly available dataset on buildings and addresses in the Netherlands.

In collaboration with [Rijkswaterstaat](#), [RIVM](#) and [Kadaster](#), Dukai's research group are reusing a collection of datasets (i.e. [BGT](#), [AHN](#) and [3D BAG](#)) to investigate how [3D geospatial data on noise sources](#) can be automatically generated for the whole of the Netherlands. This project is a great example of a relevant urban application that leverages existing geospatial data. "The development aims to improve efficiency by reusing datasets thereby eliminating the need to acquire new data," says Dukai. "Of equal importance is the opportunity to standardise input data for noise studies. The collection of 3D geospatial data by different research groups using various methods has led to undesirable differences in results which cannot be compared unambiguously."

Data harmonisation on the horizon

Dukai is an advocate for standardising open geospatial data and supporting software, and is determined to improve data quality by ensuring that it is [findable, accessible, interoperable and reusable \(FAIR\)](#). To this end, he is involved in the ongoing development of '[CityJSON](#)', a JSON-based exchange format for the [CityGML](#) data model. CityGML, or Geography Markup Language, is an international standard produced by the [Open Geospatial Consortium \(OGC\)](#) that is used to store and exchange 3D city models. However, [it's XML-based exchange format has drawbacks](#); it's complex and lacks software support. Dukai makes the case for switching to a simpler alternative, "CityJSON has been designed with programmers in mind. It's compact and easy-to-use whether you're reading or creating datasets. With one click you can [convert CityGML to JSON files](#) which are supported by a growing number of [software platforms](#)."

CityJSON's compatible software component, [cijo](#) (CityJSON input/output) is a [Python](#) command-line interface (CLI) programme designed to process, manipulate and validate CityJSON files. Since

both [CityJSON](#) and [cijo](#) are open to the vibrant research community via [GitHub](#), iterative workflows facilitate their continual improvement through the collaborative efforts of those who wish to contribute to the common goal of harmonising geospatial data. (To learn more about [#CityJSON](#), (see Dukai's [GeoPython2019 conference slides](#)).

Becoming a Geospatial Data Champion

Our conversation ended with Dukai's final thoughts on the changing landscape of open geospatial data and a positive reflection on his impact as a Data Champion. "As a software engineer, I understand the pain of having to work with proprietary data. Whilst making data 'open' doesn't necessarily guarantee quality, it does guarantee transparency and allows people to learn from one another and work together to accelerate research." He adds, "As an individual researcher, becoming a Data Champion has allowed me to acquire valuable skills and knowledge that I wouldn't have otherwise encountered. As part of a wider research community, becoming a Data Champion has allowed me to connect my group with a diverse network with opportunity for collaboration."

Mapping the future landscape...

Dukai's research group have recently been collaborating with TU Delft Associate Professor and Research Director, [Frank van der Hoeven](#); Research coordinator, [Susan Ng-A-Tham](#), and [Data Stewards](#); [Yan Wang \(Faculty of Architecture and the Built Environment\)](#), [Kees den Heijer \(Faculty of Civil Engineering and Geosciences\)](#) and [Jeff Love \(Faculty of Industrial Design and Engineering\)](#), to develop a [Geoportal](#) that will facilitate the visualisation of geospatial data via maps...Indeed, this could include geospatial data archived in our [4TU. Data repository!](#)

Carve your niche with 'The Carpentries'.





[Victor Koppejan](#) is a PhD student in computational [Bioprocess Engineering](#). His doctoral research involves the purification of proteins from biological matrices for industrial use using expanded bed adsorption. Koppejan employs open source computer simulation tools to [model the dynamics of fluid and particle flow in a fluidized bed](#) using the [Dutch National Super Computer](#).

[Raúl Ortiz Merino](#) is a Postdoctoral researcher in [Industrial Microbiology](#). He works within the field of comparative genomics to characterise microbial gene sequences of commercial relevance. After years spent training as a wet lab biochemist, Ortiz Merino made his transition to dry lab computational science and is now an experienced bioinformatician.

TU Delft Library met [Data Champions](#) from the [Department of Biotechnology](#), Victor Koppejan and Raúl A. Ortiz Merino, to celebrate their Software and Data Carpentry workshop success

Both [Data Champions](#) have made a significant contribution to their local research community by sharing their knowledge and expertise during Software and Data Carpentry workshops

What are 'The Carpentries'?

[The Carpentries](#) are a non-profit project, formed in January 2018, to teach basic computing skills to researchers worldwide. The aim is to train and foster an active, inclusive and diverse community of learners and instructors that promotes efficient, open and reproducible computational research.

During two-day Carpentry workshops, instructors and helpers share their mission to teach foundational coding and data skills using openly-available lesson material and evidence-based teaching practices. Anyone can register to attend, no matter their skill level, and what's more, at TU Delft it's free to join!

Bringing The Carpentries to TU Delft

[The 4TU.Centre for Research Data](#) became a [Gold Member](#) of The Carpentries to bring instructor training and workshops to TU Delft. Since [piloting Software Carpentry](#) for the first time in November 2019, the university has hosted two Software Carpentry workshops. After helping in the first workshop, Koppejan decided to [become a certified Carpentries instructor](#) in February and successfully held his own [Software Carpentry workshop in March](#) this year! The [most recent Software Carpentry workshop](#) took place on 8-9th July.

Ortiz Merino has volunteered his help during all workshops at TU Delft and has conducted his own whilst undertaking his certified instructor training. To build upon the experience he gained during the Software Carpentries, Ortiz Merino was financially supported by [The Data Champion travel fund](#) to join the [Introduction to Reproducible Genomics: Data Carpentry](#) in Ghent (Belgium). He used this valuable experience to collaborate with fellow Bioinformatician and Data Champion, [Marcel van den Broek](#), and organise [the first TU Delft Data Carpentry workshop](#) in June. You can read more about TU Delft's Data Carpentry workshop [here](#).



Caption: Researchers attend TU Delft's first Genomics Carpentry to 'shape up' their data science skills!

Software or Data Carpentry? What's the difference?

[Software Carpentry workshops](#) are designed for researchers who want to learn how to programme more effectively. Typically, three core topics are taught; [The Unix shell](#), [version control with Git](#), and a programming language ([Python](#) or [R](#)).

[Data Carpentry workshops](#) are designed for researchers who are dealing with domain-specific data. The workshops are centred around a single dataset and teach participants [project organisation and management](#), [introduction to the command line](#), [data wrangling and processing](#), and introduction to [cloud computing for genomics](#).

Why invest in The Carpentries?

We heard why both Data Champions elect to use Carpentry workshops as a means of disseminating knowledge amongst a wider audience.

"I became inspired by [open online training on high performance computing](#) provided by [Argonne National Laboratory \(USA\)](#)," says Koppejan. "After undertaking online tutorials, I was enthusiastic to share the knowledge I'd gained with my colleagues but didn't have sufficient time to train them all on a one-to-one basis." He continues, "I believed that becoming a Carpentries instructor would help me spread the word of good code management amongst a larger research community."

Ortiz Merino shared similar motivations. His research section comprises 4 principal investigators, 4 postdoctoral researchers, 16 PhD students and 10 technicians, not to mention the constant flux of Masters and Bachelors degree students that can reach as many as 50 individuals. He also uses Carpentry workshops to reach more people. "Most members of my section encounter similar research problems I thought The Carpentries workshops would make it easier to gather together to answer queries, explain common concepts and learn as a group."

Carve your niche, be your own bioinformatician

Bench scientist turned computational biologist, Ortiz Merino, understands the challenges of moving from the wet to dry lab environment. “Nowadays, it’s difficult for biologists to avoid computational approaches all together. Most modern scientists will have to learn computer programming at some point during their career.” He reflects on his personal experience. “Making the switch is not easy. It took me several years to learn the specialist data science skills required to make my transition from experimental to computational biology. Working as an intermediary between the two spheres, I want to bring the wet and dry lab closer together and I believe Data Carpentry workshops can help me to achieve this.”

The workshops introduce wet lab scientists to computational tools in an approachable way, bridging the gap between generating and analysing data. “Participants receive all of the basic information they need in a structured two-day workshop so that they can start learning how to become their own bioinformatician.” Ortiz Merino assures that Data Carpentry workshops are the best way to learn.

Sculpt your Soft Skills

Aside teaching technical skills, The Carpentries teach soft skills that enhance personal and professional development. Koppejan explains how training to become a certified instructor aided in the development of his interpersonal skills. “[The instructor training programme](#) taught me how to communicate more effectively and interact harmoniously with workshop participants. I became more conscious of listening and teaching with empathy.”

Koppejan emphasised the importance of creating an inclusive, interactive and collaborative learning environment. He recounted his positive experience of attending the [Collaborations Workshop 2019 \(CW19\)](#), an ‘un-conference’ that brought multidisciplinary personnel together to Loughborough University (UK), to explore best practices and the future of research software in a relaxed, social setting.

[The rules of an ‘un-conference’](#) are simple:

- #1. Whoever shows up are the right people.
- #2. Whatever happens is fine.
- #3. Whenever it starts is the right time.
- #4. It's over when it's over.

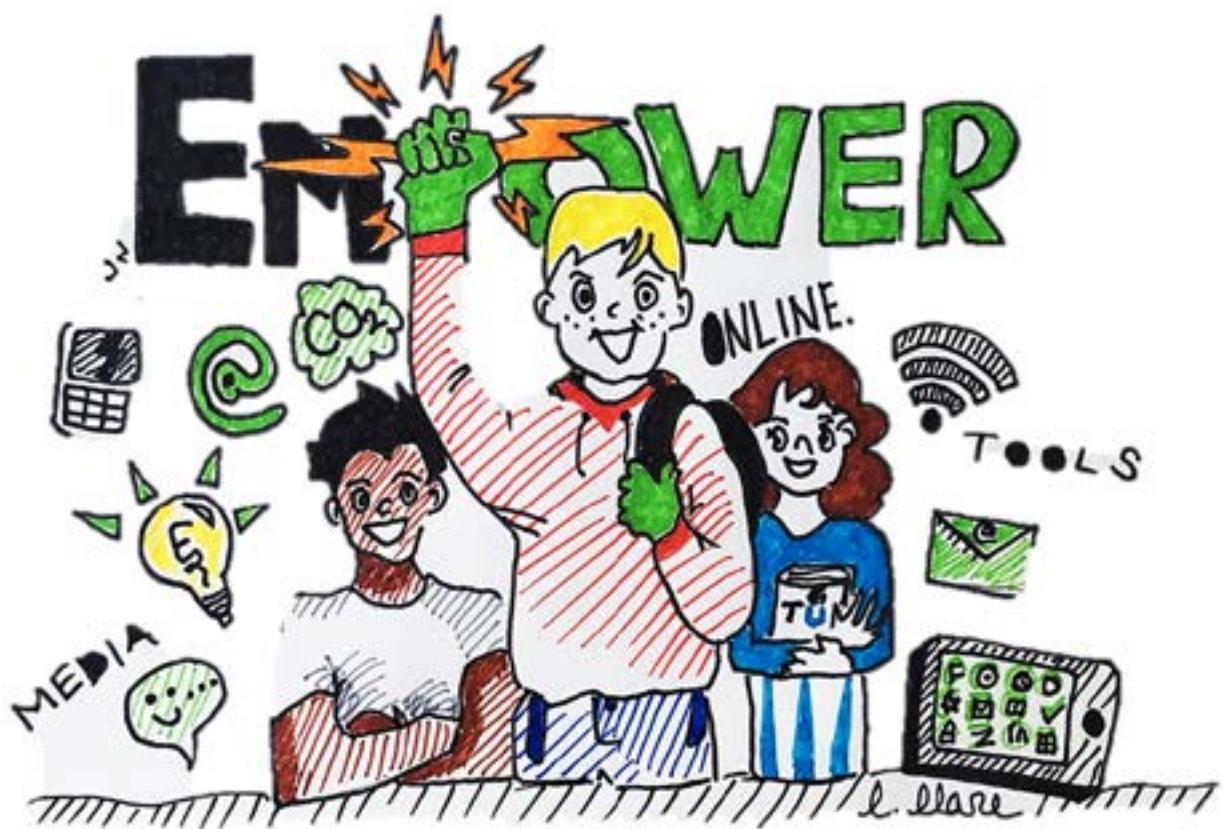
And, not forgetting ‘The Pac-Man Rule’: When standing in a social circle, always leave enough space to encourage new people to join the group conversation!

We wish our Data Champions good luck in their bright future!

Koppejan’s personal attributes, and the transferable skills he has developed during his time as a Data Champion and Carpentries instructor at TU Delft, have led to exciting career prospects. It’s pleasing to hear that after expressing his interest in Open and FAIR data in a recent job interview, he has secured a position as a Data Scientist at DSM and will start his new role in October!

We send our grateful thanks and best wishes to Victor Koppejan as he enters the next chapter of his career, and also to Raúl Ortiz Merino as he becomes a certified Carpentries instructor at TU Delft.

Empowering people with data.





Data Champion from the [Faculty of Industrial Design and Engineering](#), [Natalia Romero Herrera](#), explores innovative ways of giving data back to the user for social good.

Chilean computer scientist, [Natalia Romero Herrera](#), has worked in the field of human-computer interaction for almost 20 years. Her research within the [Design Conceptualisation and Communication](#) section at TU Delft focuses on the development and application of user-centred design methods to understand people's daily life practices and, in particular, the impact of environmental and health issues.

As an active member of the [idStudioLab](#), Romero Herrera uses [living labs](#) to contextualise complex behavioural data. In a living lab, innovative technological applications and tools are designed, and tested by participants to determine how people engage with specific products and services. Romero Herrera's technological designs are 'experience-centred', meaning that they aim to improve user experience as a whole. Therefore, a key element of her research is to learn about the needs and values of users; to engage, empathise and empower them through data.

A 'give and take' perspective on data

"Technology has always used data to direct people towards a desired goal," explains Romero Herrera. "For example, environmental technologies use data to encourage people to reduce their energy consumption at home. Likewise, healthcare technologies use data to encourage people to walk 10,000 steps a day because it's good for their health."

Whilst data plays a central role to drive change in societal behaviour, Romero Herrera has different ideas for its use. "Rather than simply taking data from people, we can give data back to people to influence their actions." She enlightens us on her mission to empower citizens and communities by giving them data. "My research aims to grant people access to the information they need to make autonomous decisions and change their own behaviour."

To this effect, we discussed Romero Herrera's current international projects that give data back to people for personal and wider social benefit.

Getting students involved with 'ENERGE'

The EU has set itself the goal of [reducing greenhouse gas \(GHG\) emissions](#) to zero by 2050. [ENERGE](#) (ENergizing Education to Reduce Greenhouse Gas Emissions) is a project funded by [InterregNWE](#) that helps to achieve this goal by facilitating the implementation of low carbon, energy and climate protection strategies to reduce GHG emissions in North-Western Europe.

The project focuses on secondary schools which are typically housed in pre-war buildings and are, therefore, energy inefficient. "If such schools are to meet the zero emissions target, a significant financial investment will be required to renovate their infrastructure." Romero Herrera continues, "As these renovations will take a long time to plan, there is a demand for immediate low-cost solutions that enable long-term resource efficiency and reduced GHG emissions."

ENERGE aims to achieve a 15% reduction in total energy consumption within 12 demonstration site schools in the UK, Ireland, France, Germany, Luxembourg and the Netherlands over a four-year period. Here's how...

Quantity and quality with mixed methods

"We're collaborating with research institutes and industry partners to design mixed method tools to capture quantitative and qualitative data from schools. This will involve targeted physical interventions, such as a web-based platform, energy meters and building sensors that measure temperature, humidity, carbon dioxide, sound and light." Romero Herrera has used sensors to collect indoor climate data from buildings in past research projects.

Objective (quantitative) data collected from sensors will be combined with subjective (qualitative) data collected from behavioural studies involving students. Using similar tools to those developed for her previous [SusLabNWE](#) project, Romero Herrera can learn more about the comfort of the classroom by asking students how they are feeling. For instance, students can report their thermal comfort using a dial. A digital diary presents the data from the dial in a daily timeline and invites users to indicate what they did to manage their comfort... Did they have to open a window? Or, turn on the heater?

Education to empower

Ultimately, the data collected from ENERGE is given back to the students, along with various data enabled tools, workshops and a hackathon. Students are educated on the data lifecycle as they gain first-hand experience of data capture, processing, analysis and visualisation.

Romero Herrera adds that students also learn the importance of collaboration. "Taking a holistic and multidisciplinary approach involves many stakeholders within the school ecosystem. Students, teachers and



Caption: Mixed method tools used in Romero Herrera's research to collect integrated indoor climate data. (SusLabNWE and Building Occupant Certificate System (BOCS) Climate-KIC).

managers collectively engage with data, learn about the impacts of their actions on energy consumption and can experiment to develop new strategies to solve problems together.”

She also highlights an important future impact of ENERGE. “The project raises awareness about environmental sustainability amongst our future generation. It gives students agency to reduce their energy consumption and help to mitigate GHG emissions. These teenagers are our planet’s future decision-makers and it gives them a voice to express their ideas and opinions.”

The long-term impacts of ENERGE will be consolidated by the addition of revised educational material to supplement existing school curricula for secondary school students. Moreover, the project will monitor the effects of its initiatives beyond the school environment. Data captured in staff and student homes will emphasise the importance of sustainable energy efficiency within the domestic environment.

A healthy relationship with data

[Food Sampler](#) is another of Romero Herrera’s current projects. Funded by the [ZonMw Create Health](#) programme, this project uses mixed methods to monitor food intake in overweight or obese adult patients.

Romero Herrera describes the problems of existing methods for monitoring food intakes. “Paper-based questionnaires are laborious, time-consuming and are often not applicable to real life. Patients are, therefore, reluctant to complete questionnaires.” She advises, “If we are to be successful in changing food consumption habits of overweight individuals, we must develop better reporting techniques that engage participation and extend data collection with contextual aspects of dietary practice.”

Persuasive engagement with patients

Food Sampler integrates objective data from tools and subjective data from patient self-reports to evaluate complex dietary behaviour.



Upon interviewing patients, Romero Herrera found that their main objection to completing questionnaires was not due to labour or time, but their fear of judgement. Design has a main role in redefining the qualities of current reporting practices to reduce these negative experiences.

She explains the importance of designing non-intrusive in-situ mixed methods to persuade patients to engage with reporting tools. "Future methods must identify not what individuals are overeating but why they are overeating in order to understand the ecology of food intake."

In the project's preliminary stages, Romero Herrera is trialling several prototypes in living labs. "Inspired by the generation of e-health prevention apps, I'm developing ideas for tools that can be used in people's homes. I want to design a tool with a user-friendly interface that encourages patients to confide. Like a secret diary, patients can report on their mood, emotion and other contextual factors that influence their food intake in a way that is non-judgemental and non-confrontational."

As with all of Romero Herrera's research projects, the resulting data will be given back to patients in a way that they can understand and relate. She hopes that Food Sampler will give patients the relevant knowledge to reflect and take action to improve their personal health and wellbeing.

Publishing personal data

Romero Herrera discusses her ambitions to openly publish her datasets. "My datasets are rich. The data is not only relevant to designers but many other professional fields. I'm interested to see how making my data accessible and referable can benefit other scientific research communities."

She talks about the challenges of sharing personal data. "Of course, the data I have collected during my research is confidential and extremely sensitive. Therefore, if it is made public, it must be anonymised."

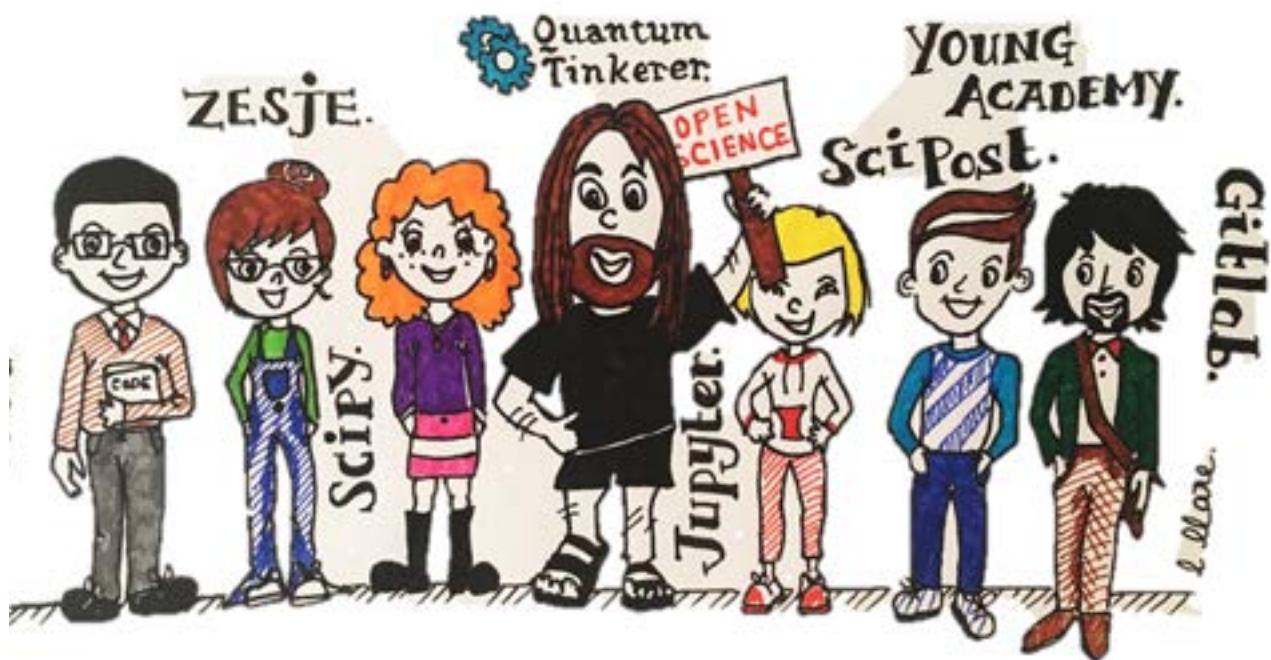
After guidance and support from faculty Data Steward, [Jeff Love](#), Romero Herrera has investigated ways to anonymise her data by clustering patients into categories so that individuals cannot be identified. She is now confident to deposit a coded version of her Food Sampler dataset on the [4TU.ResearchData repository](#).

Back to Chile!

As an editor-in-chief of the open access journal, [EAI Endorsed Transactions on Pervasive Health and Technology](#), Romero Herrera, shares her interest of empowering people through data with the wider scientific community. Next year she will chair the first edition of the EAI International Conference on Digitalising Healthcare, '[DigiCare 2020](#)', in Santiago (Chile), inviting researchers, designers, developers and policy makers to further explore the role of technological innovation in solving societal health challenges.

Thank you Natalia for showing us a new side to citizen science... “Rather than simply taking data from people, we can give data back to people to influence their actions.”

Starting small and thinking big: A Quantum Tinkerer's quest to mentor Open Science.





We hear from [Data Champion](#) and Associate Professor in [Quantum Nanoscience](#), [Anton Akhmerov](#), about his journey towards becoming an ambassador for Open Science at TU Delft.

An individual who volunteers their discipline-specific expertise, promotes FAIR data principles and advocates proper research data management.

An individual who uses their passion for knowledge exchange and their desire to build a collaborative and researcher-led community to drive the uptake of Open Science within their faculties and departments.

By definition, [Anton Akhmerov](#) leads by example as an outstanding TU Delft Data Champion.

Inspired by open-source

[Akhmerov](#) is a theoretical physicist at the [Kavli Institute of Nanoscience](#) and [QuTech](#) at TU Delft. His expertise in quantum mechanics and nano-superconductors involves numerical simulations to explore complex phenomena, such as topological quantum computation. His research uses various open-source software, including [SciPy](#) (a Python-based software for STEM subjects) and [Jupyter](#) (a project that exists to develop open-source software, standards and services for interactive computing).

Akhmerov explains why open-source software is integral to his work. “Many of my ideas aren’t radically new; I see something good and I want to make it even more awesome.” He adds, “Sharing code increases the impact of such work, and makes it possible to work on topics as a community.” To this end, we learned about Akhmerov’s contribution to Jupyter whereby he created a library for [incorporating code output in documentation using Jupyter-Sphinx](#) together with collaborators. His example demonstrates how incremental improvements can take place through open-source. He also feels compelled to make his research open for altruistic reasons. “Since my software development is achieved within the open-source domain, it seems natural to share my code so that others can use it.”

Starting small

Starting small with his ambitions, Akhmerov began by perpetuating his culture of sharing amongst colleagues of the [Quantum Tinkerer](#) research group at TU Delft. After back-and-forth discussion about how they could make their research more open, it was agreed that the group's work in progress would be shared internally on their [Kwant Gitlab server](#) whilst finalised code would be published openly on [zenodo](#).



Caption: Anton Akhmerov and his colleagues of the Quantum Tinkerer research group at TU Delft

Thinking big

In collaboration with another Data Champion [Gary Steele](#), Akhmerov worked to develop and implement an [Open Data Policy](#) to engender a culture change across the entire Quantum Nanoscience Department. The policy provides guidelines on preparing and openly publishing data for the department but is to be adopted across the [Faculty of Applied Sciences](#) in the hope that it will eventually become the norm at TU Delft.

He also actively engages with the wider scientific community to stimulate a discourse about making research data and software open. His talk, '[Time to share is now](#)', was delivered at the [Data Champions kick off meeting](#) in December 2018 to emphasise the importance of software in contemporary research. He was also an author of the presentation, '[Making Research Software a First-Class Citizen in Research](#)', that was delivered at a meeting with The Netherlands Organisation for Scientific Research (NWO) in March. The presentation advances the idea that if we are to produce transparent, reliable and reproducible research in the name of Open Science, then publications, data and software must be treated on equal footing at the policy level.

The Spirit of Secrecy

Encouraging researchers to make their software open isn't easy. "Many researchers see sharing code as a radical idea and may object for several reasons; they may feel like their code is invaluable to others or that it takes too much time to clean up and make readable. Alternatively, they may have dedicated a lot of time to develop a piece of code that gives them a competitive advantage and, therefore, feel reluctant to share it." Says Akhmerov.

Whilst he understands such objections, his personal perspective is somewhat different. “I don’t particularly appreciate this spirit of secrecy and I hope the prominence of this viewpoint will eventually diminish.” He claims, “Researchers who want to publish their papers but refuse to publish their code aren’t helping to solve the problem of reproducibility.”

Reviewers, rethink! Do something for Open Science that requires zero effort.

Akhmerov realised that he could do something to help solve the problem with no extra effort. As a reviewer, he began requesting all of the underlying data and code for each paper that he reviews. “It’s the responsibility of the reviewer to ensure that the publication is based on sound results. By examining the data and code used in the paper, one can verify whether or not the findings are indeed true.” His article, '[I need your data, your code and your DOI](#)', relays the message that reviewers are in a position of power and can mentor Open Science by ensuring quality standards are met.

He is a member of the editorial board of the [New Journal of Physics](#), and more recently, [SciPost](#); an online, open access, community-run journal managed by scientists. Akhmerov highlights the benefits of SciPost’s guiding principles and is in favour of making the peer-review process more visible to the scientific community. “SciPost assigns each manuscript review a DOI, meaning that the review process can be witnessed and is no longer confidential. This leads to a stricter peer-review process as reviewers are held accountable. What’s more, reviewers can be credited for their time and effort invested in evaluating submissions.” Using publication portals where public-funded science is freely and openly accessible to anyone is an effective way to advance Open Science worldwide.

Open Education

Akhmerov leads by example as a mentor for Open Science. He teaches the undergraduate Solid State Physics course at TU Delft and has created publicly open [online lectures notes](#) with simulations. The source code is openly available so that students can use and modify the lecture material. He explains how this aids the learning process, “Students are encouraged to engage with the lecture material. Since they can access the code, they can correct it, develop it and can even design their own lectures.” Akhmerov’s lectures are a great example of self-directed, researcher-led learning. Moreover, making lecture material open means that it can be used by many course lecturers simultaneously, making teaching more efficient. To this effect, he also conducted a [Massive Open Online Course \(MOOC\)](#) to educate thousands of students about topology in quantum mechanics.

In collaboration with software developer, [Joseph Weston](#), Akhmerov developed '[Zesje](#)', a web app where written exam manuscripts are scanned and systematically graded on a question-by-question basis. The app, named after the Dutch term ‘zesjecultuur’ which means ‘Grade C-’, was devised to streamline the grading process by assessing exam manuscripts electronically, rather than on paper. Akhmerov came up with the idea when faced with the daunting task of marking around 300 physics undergraduate exam papers. “Whilst an examiner may have a predefined idea of how they want to grade exam answers, it’s extremely difficult to ensure consistency when assessing a large cohort of students on paper.” He adds, “Zesje saves time and

effort. An examiner can mark more consistently and the grading can be distributed throughout the entire course team.” Akhmerov’s enthusiasm for [low effort exam](#) grading has spread across the university. Now, around 20 courses use Zesje, including those within the Computer Science department that comprises approximately 900 students. Akhmerov was awarded the [2018 Delft Educational Fellowship](#) for his innovation.

Dedicated to advocating good practice in open-source software development, Akhmerov organised a [one week course](#), inspired by [Software Carpentry](#), to teach basic programming skills to PhD students of the Casimir Research School.

Future ideas

As a member of the [Young Academy](#), he is part of a dynamic team of scientists who share a broad interest in science practice, policy and communication. His group project within the academy aims to reduce the carbon footprint of academic travel through the organisation of [virtual conferences](#). “Aside reducing carbon footprint, online conferences have several advantages over offline conferences,” says Akhmerov. “They reach a broader audience and are more inclusive since they accommodate individuals who are limited by funding or prohibitive travel logistics. Online conferences are also less administrative, easier to organise and cheaper to host.” We look forward to hearing how his idea to run virtual conferences at TU Delft progresses.

We are truly inspired by the positive energy and enthusiasm Akhmerov brings to TU Delft as a Data Champion and mentor for Open Science. Learn more about his latest ideas and perspectives by following his [blog](#) and twitter accounts; [@AkhmerovAnton](#) and [@QuantumTinkerer](#), to keep-up-to-date!

Five W's of Open Data.





TU Delft library met Open Data Expert and [Data Champion, Anneke Zuiderwijk](#), to learn about the ‘who, what, where, when and why’ of Open Data.

Anneke Zuiderwijk is a [Data Champion](#) working within the [Faculty of Technology, Policy and Management](#). As an Assistant Professor in the Department of [Engineering, Systems and Services](#), she has dedicated her academic career to investigating the complex rationale behind open data. In Zuiderwijk’s unique case, data sharing isn’t just a research consideration, rather it’s her research domain.

Zuiderwijk believes that great value can be created through open data sharing and use and that it can be revolutionary for scientific advancement. Her mission is to develop theory for designing infrastructures and institutional arrangements that incentivize open data sharing so that it eventually becomes the standard rather than the exception.

In order to realise her mission, her academic research and responsibilities as a Data Champion address the following ‘5W’s of open data’:

1. Who are the actors involved?
2. What are their motivations to share data? And, what can infrastructure do to motivate data sharing?
3. Where can data be shared?
4. When should they prepare to share data?
5. Why consider data sharing?

About Anneke

Before joining TU Delft in 2011, Zuiderwijk worked for the [Dutch Ministry of Justice and Security](#) where she gleaned insight into the real world of open judicial data. “The Ministry were collecting data on crime statistics and were faced with many challenges of sharing such sensitive information with the wider community.”

Motivated to mitigate the challenges associated with sharing and using government data, Zuiderwijk began her PhD under the supervision of [Professor Marijn Janssen](#) at TU Delft. Her [doctoral research](#) involved the design of a socio-technical infrastructure that enhanced the coordination of open government data use. By creating a metadata model with interaction and data quality mechanisms, Zuiderwijk aimed to stimulate interaction between data providers, data users and policy makers thereby creating a feedback loop to better inform policy about government data use.

After obtaining her PhD with distinction, Zuiderwijk assumed a postdoctoral researcher position working on the [VRE4EIC](#) project, funded by the European Union's 3-year Horizon 2020 research and innovation programme. 'VRE4EIC' stands for 'Virtual Research Environment to Empower multidisciplinary research communities and accelerate Innovation and Collaboration'. Essentially, this online platform draws professionals from various scientific domains to exchange data, resources and knowledge within a multidisciplinary data environment. Data can be made open at different levels within the platform meaning that there is control over data sharing and data can be protected if necessary.

Zuiderwijk emphasises the importance of embracing diversity and fostering inclusion for scientific progression. "Many research environments focus on discrete scientific disciplines with limited convergence, yet, as scientific problems are related across all disciplines, it's of tremendous value to bring disciplines together to solve problems collectively. After all, societal problems must be studied from multiple perspectives."

1. Who are the actors involved?

The field of open data involves a number of actors. "Researchers, governments, companies, citizens, journalists, librarians and archivists are all concerned with open data to some extent," informs Zuiderwijk. A key objective of her research is to understand the needs of these different actors and how to motivate them to openly share their data and use the data of others. "In order to build a sociotechnical infrastructure that facilitates effective data sharing across multiple disciplines, it's necessary to understand the perspectives, behaviours and motivations of each actor involved."

2. What are their motivations to share data? And, what can infrastructure do to motivate data sharing?

She explains that their motivations are diverse and discipline-specific. "Researchers may be motivated to openly share data if it means greater visibility and more citations to boost their career development." However, Zuiderwijk admits, "There are still obstacles that prevent researchers from sharing data, such as a lack of the required skills."

Whilst the majority of obstacles cannot be mitigated completely, Zuiderwijk explains that implementing suitable infrastructure and related institutional arrangements can help. "It's important to improve the ease of use of data infrastructures, of which open data portals are one element, and to provide institutional training for using such portals so that researchers acquire the necessary skills and become motivated to share data."

Data sharing is more common in some academic disciplines than others and there are obvious disparities across them. Her recent publication, '[Sharing and re-using open data: A case study of motivations in astrophysics](#)', explores the complex interaction of factors influencing open data sharing and use in astrophysics, a single scientific discipline where data is already extensively shared and reused. Her case study demonstrates that benefits associated with data sharing involve intrinsic motivations, such as more reproducible science and faster rates of scientific advancement. Zuiderwijk hopes that insights gained from her study of astrophysics can be transferred to disciplines where data sharing is less common.

She discusses the motivations of other actors. "Whilst governmental organisations are typically motivated to share data for public value, it's more difficult for corporate organisations to do so." She continues, "Commercial enterprises are often restricted by proprietary interests, trade secrets, and concerns over legal ramifications of privacy and security breaches." Zuiderwijk is ambitious to help overcome the barriers of data sharing and use within the private sector and build trust through the creation of mechanisms for sharing open business data. This is an objective she wishes to explore in the near future.

3. Where can data be shared?

Data is typically shared via online platforms. Zuiderwijk uses the [4TU.ResearchData](#) repository to openly publish her data. She believes that one of her most important contributions as a Data Champion is motivating others to follow her practice. "I actively encourage my students to publish their data on the 4TU repository." She adds, "Working transparently gives them more credibility as early career researchers. If a reviewer can see that their underlying data is openly accessible it instils confidence that the findings presented within their manuscript are robust and reliable."

4. When should they prepare to share data?

As soon as Zuiderwijk's students begin their research, she educates them on the principles of Open Science and proper research data management. She ensures that they each receive relevant training to openly publish their data upon completion of their research project. "From the point of data collection, students are trained on how to cleanse and curate their data so that it's understandable to others when they openly share it. Formatting data correctly saves time and effort when it comes to uploading data to the repository."

Zuiderwijk coordinates with faculty [Data Steward, Nicolas Dintzner](#), to help students create their data management plans prior to conducting research. Wherever necessary, she also mandates that students apply to the research ethics committee so that they understand the regulations and requirements of collecting and sharing confidential data.

"Reluctance to share confidential data often arises because researchers believe that data anonymisation costs extra time and effort." She argues against this belief, using her own research methodology as an example. "When I systematically analyse personal data, such as interview transcripts, I assign codes to the data to describe the content. This means the codebook; an Excel spreadsheet containing analysed data, is already anonymised. With no extra labour, the coded version can be openly published on the repository to benefit the wider community."

5. Why consider data sharing?

Despite important considerations for confidentiality, there's no doubt that the arguments for sharing data are powerful. Individuals have an opportunity to advance their scientific knowledge, and improve their visibility and credibility, which in turn raises their professional profile. They have more scope to build international, multidisciplinary connections and foster a collaborative community wherein resources can be shared for social and economic gain.

Whilst the benefits are obvious, Zuiderwijk admits that the culture change towards open data is slow. "It's a long-term transition. It took me several years to appreciate why I should share my data," she says. "I'm fortunate that in my field of research I get to experience the benefits of data sharing and reuse. Unfortunately, actors in other disciplines rarely get to witness the advantages."

Passionate about educating others about the importance of data sharing, Zuiderwijk has instructed the ProfEd course on [Open Data Governance: From Policy and Use](#) and the Massive Open Online Course (MOOC), [Open Science: Sharing Your Research with the World](#), together with TU Delft Library instructors, [Michiel de Jong](#) and [Nicole Will](#). Alongside the leader of TU Delft's [Knowledge Centre Open Data](#) (Kenniscentrum Open Data), [Bastiaan van Loenen](#), she will lecture during the upcoming [MOOC on Open Government](#) organised by [Marijn Janssen](#), which takes place in September this year.

Switch gear! Drive the uptake of Open Science within your research team.





Associate Professor, [Joost de Winter](#), and Postdoctoral Researcher, [Pavlo Bazilinsky](#), share their experience of working together on automated driving projects

Associate Professor, [Joost de Winter](#), works in the [Department of Cognitive Robotics](#). His research team investigates human factors of automated driving, for example, how drivers, pedestrians, cyclists and other road users interact with highly automated vehicles. Their aim is to develop automated vehicles that can understand and adapt to the driver's state and intentions for safer transportation. de Winter advocates Open Science and transparent working within his team to improve the validity and reproducibility of their research.

Postdoctoral researcher, [Pavlo Bazilinsky](#), has been a valued member of de Winter's team since 2014 when he began his PhD. His doctoral research explored the use of auditory interfaces for human interaction with automated vehicles. Bazilinsky received several prestigious awards for his doctoral work, including the European Young Researchers Award, MCAA Best Innovator Award and EMA Alumni Award. His current research focuses on how automated vehicles should communicate with human road users.

Both scientists recently joined our vibrant team of [Data Champions](#) to drive the uptake of Open Science amongst the research community of TU Delft.

A new avenue

"It was around five years ago when I began to explore the field of Open Science," says de Winter. "I became concerned about the reproducibility crisis which is a serious problem in Psychology and other scientific disciplines." He continues, "I realised the need to be more careful about conducting empirical research and believed that the scientific process could be made more open to improve the reproducibility and credibility of published results."

With this in mind, de Winter has taken positive steps to promote Open Science within his own research team and is enthusiastic to share his wisdom with a wider scientific audience. His recent talk at the [Future Forward seminar](#) emphasises the importance of replication studies and offers valuable advice on how researchers can achieve reproducibility. For a summary of his talk, read '[Can I really read your emotions if I look deep into your eyes?](#)'

The 'open' road

With a team of 6 PhD students, 8 Master's students and a cohort of Bachelor's degree students, de Winter understands the value of establishing proper research data management infrastructure for efficient coordination and collaboration. "Since I manage several large research projects which overlap in content, there are usually a number of researchers working together on the same dataset. Therefore, it's imperative that all team members have access to the data repository and that they know how to store and archive their data for effective data sharing." He highlights that this is particularly important for the seamless continuation of research projects when there is a frequent turnover of students.



As part of their data management plan, de Winter obligates that his students adhere to a specific project folder and file structure that can be openly accessed via [Dropbox](#) or [Google Drive](#) from the start of their project. Postdoctoral researcher, Bazilinsky, demonstrates how students can organise their work by sharing examples. "I give students access to my past projects so that they learn how to appropriately format and manage their own data." He also introduced [Slack](#), a communication tool that facilitates knowledge exchange amongst team members.

de Winter expects his students to publish their experimental data on the [4TU.ResearchData](#) repository upon project completion. "It's not enough to publish results tables and figures in a publication, researchers must also publish their underlying data." He explains that making data open holds researchers accountable for their work and helps to solve the reproducibility crisis. "Publishing data adds value to a scientific publication. If anyone has questions about the results of an experiment they can examine the raw data, validate it and reuse it." de Winter hopes that open data practice will spread beyond his research group and gain traction within the entire [research faculty](#).

In the driving seat

"Young scientists are the answer to driving a culture change," says de Winter. "New students embrace Open Science. They aren't afraid of adopting contemporary standards; they're proud to take ownership of their independent project and see the value of working openly."

Bazilinskyy is a good example of a modern scientist with an intrinsic desire to drive Open Science. He explains that his passion stems from his long-standing appreciation for open-source software. “Ever since undertaking my Bachelor’s degree in Computer Science, I’ve believed that science should be open. I want to collaborate with others and make my data available to advance scientific progress.”

Accelerating science

As a Marie Curie fellow, Bazilinskyy worked on the [Human Factors of Automated Driving \(HF Auto\)](#) work package, [‘Human-machine interface of the future highly automated vehicle’](#). He explains the purpose of his PhD research. “When an automated vehicle reaches its operational limits, it must provide a takeover request to alert the driver to resume control. Current human-machine interfaces within automated vehicles provide the driver with instant auditory cues, such as warning sounds or spoken messages. Such binary cues can be problematic, leading to false alarms, distraction, information overload, and can be annoying for the driver.”

Bazilinskyy investigated the use of directional auditory feedback via speakers in the cabin of an automated truck to improve communication with the driver. Using driver simulation experiments, the use of sound was validated in isolation and in a multimodal setting. Hence, Bazilinskyy collaborated with [Alexander Eriksson](#) from the [University of Southampton](#) and [Veronika Petrovych](#) from the [Swedish National Road and Transport Institute \(VTI\)](#) to implement advanced visual cues, and [Bastiaan Petermeijer](#) from the [Technical University of Munich](#) to implement tactile cues via active inceptors, such as joysticks, the driver’s seat, steering wheel and pedals.

Based on the premise that binary auditory cues are perceived as an annoyance for drivers, Bazilinskyy collaborated with [Volvo Trucks](#) in Gothenburg, Sweden, to conduct a naturalistic driving study. In this study continuous auditory feedback was used to inform the driver about various motoring parameters, such as cruise control status, lane deviation and headway time. The continuous auditory cues blend with natural sounds of the cabin (i.e. engine and wind sounds) and keep the driver informed on a continual basis.

Backseat drivers!

As our roads become more populated with automated vehicles, there is concern for the lack of interaction between automated vehicles and human road users. “In current motoring practice it’s common for drivers to communicate with pedestrians and other road users via eye contact or hand gestures. However, in the future of automated driving, drivers will become passengers and are likely to spend their daily commute preoccupied with work, reading, playing games, or even sleeping.”

In order to determine how automated vehicles should interact with other road users, Bazilinskyy devises novel communication methods and conducts crowdsourcing experiments using the [Figure Eight Platform](#) to evaluate the efficacy of each method.

An ongoing crowdsourcing experiment uses animations to present different ways to visually communicate the intent of an oncoming automated vehicle. Respondents are asked to watch a series of animations and provide feedback about the effectiveness of each visual stimulus.



Caption: Various ways to display the intent of an oncoming vehicle through the use of visual signs

Bazilinskyy advocates the use of crowdsourcing for Open Science. “For research to be robust, reliable and reproducible, it’s important to use a large and representative sample population. Crowdsourcing draws thousands of respondents from all corners of the globe to give you more powerful scientific results.”

SD-Insights

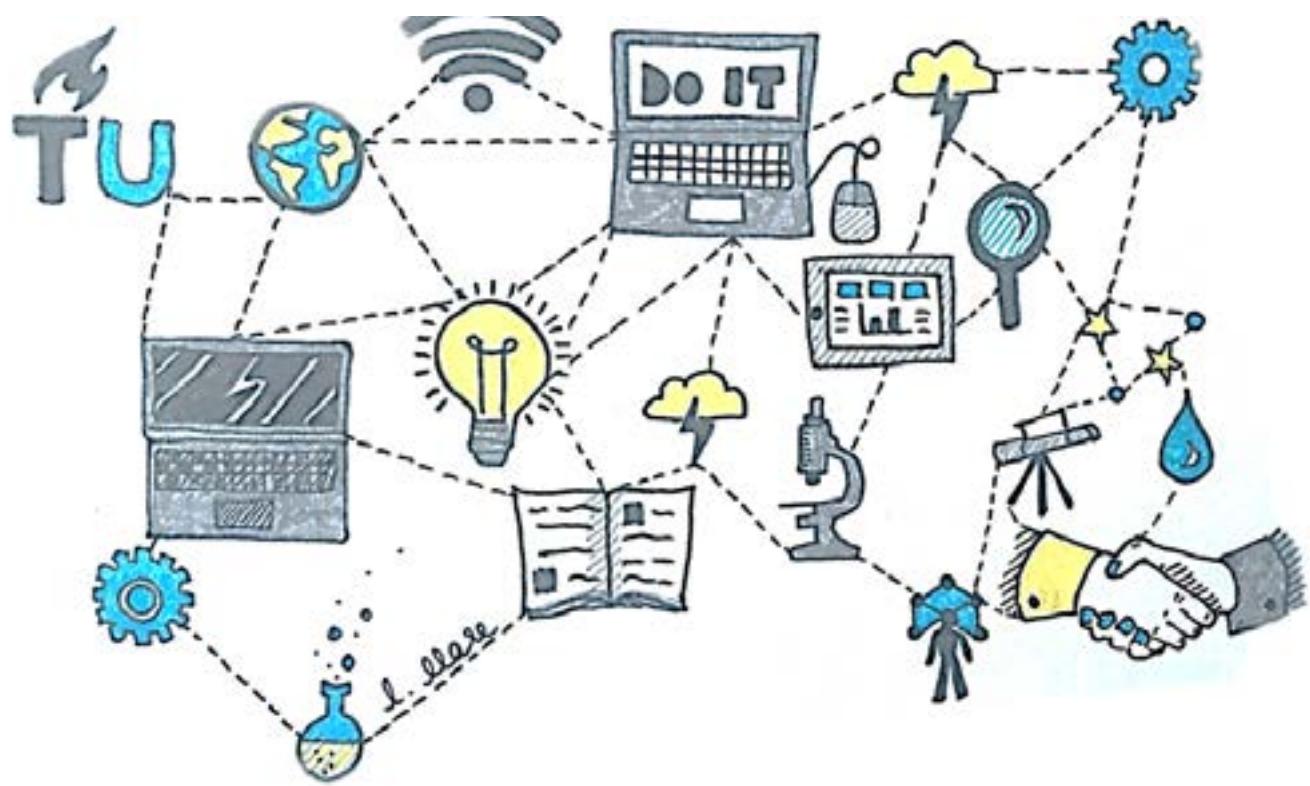
Aside his academic career, Bazilinskyy is the Director of Data Science Research at [SD-Insights](#), a corporate organisation that aims to improve driver performance, safety and efficiency.

Their main clientele are transportation companies that register their fleet of vehicles for the installation of [mobileeye](#) advanced driver assistance systems (e.g. cameras, sensors and tracking devices). Data is collected on approximately 200 driving parameters for each vehicle. Bazilinskyy analyses the data and provides drivers with personalised feedback to help them drive safely.

He believes that SD-Insights can improve profitability for transport companies who register their drivers. “Safe driving reduces company costs through better fuel economy, reduced vehicle maintenance and, ultimately, less road traffic accidents.”

This is a positive example of how data can be shared with drivers to educate and motivate them to adopt safe and efficient driving practice.

Helping researchers do IT.





"In the era of big data more researchers are dependent on digital technology. I want to help researchers find the appropriate tool for the job," says [Data Champion, Susan Branchett](#)

[Susan Branchett](#) has worked at the interface between science and software for several years. Since obtaining her PhD in Computational Quantum Mechanics at [University College London](#), she has assumed various professional positions to provide IT services and software support. Following her post as Director of Business Development and Operations at the [Netherlands eScience Center](#), she joined TU Delft's [IT Innovation Department](#) two years ago to share her expertise in Research Data Innovation.

Demand for digitalisation

Branchett aims to help researchers use IT. "In the era of big data researchers are becoming more dependent on software, algorithms and complex data analysis. I want to help them find the appropriate tool for the job." She explains her mission to become a point of contact for researchers at TU Delft. "I wish to bridge the gap between science and IT. Researchers can come to me for help and support, and I can show them how to apply IT tools to save time and improve their research."

She identifies challenges of finding suitable tools during a time of rapidly evolving IT development. "Researchers are often immersed in their work and don't have the time, know-how or inclination to keep abreast of emerging technologies. What's more, they often don't know where to seek IT support." To this end, Branchett decided to become a [Data Champion](#) to engage with the research community and raise awareness of the support services provided by the IT Department.

A survey about storage

Before you can provide researchers with IT support, you need to find out what support they require. Earlier this year, Branchett and colleagues from the IT Department disseminated a survey to learn about the storage requirements of academic staff at TU Delft.

The survey asked questions about their current needs and preferences for data storage. With 625 respondents, the results gave in-depth insight into data storage practice across all eight university faculties. “We discovered that around 75% of academics share data with students; more than 75% share data with external collaborators; and, more than 50% share data globally.” Branchett continues, “The majority of academics share confidential data with international collaborators and are using their own personal devices to do so. From this, we identify that improvements to data storage can be made.”

In response, Branchett and co-workers, [Bert Kuipers](#) and [Wim Penninx](#), submitted a report to the [Paul Hillman](#) (Head of IT) and [Bart Feenstra](#) (Manager of the Shared Service Centre) to summarise the results and propose recommendations for improving existing data storage services. Branchett [presented an overview](#) of the [survey results](#) during the latest [Data Champions](#) meeting.



Education for engagement

Working within the Innovation Department, Branchett is continually exploring new ways to support academics in their education. Last year, she organised a [Deep Learning workshop](#) which brought together 80 participants to learn basic concepts of machine learning. The workshop proved so popular that places were filled within 24 hours of its announcement! [Head of the Computer Vision Lab, Jan van Gemert](#), introduced [basic theory of machine learning](#), including the concept of ‘feed-forward’ and convolutional networks. Followed by [Osman Kayhan](#) and Data Champion, [Paul van Gent](#), who shared hands-on examples for participants to put theory into practice and experiment with their newly acquired programming skills.

In July, Branchett instructed a [Software Carpentry workshop](#) within a week of becoming certified. She reflects on her experience of delivering her first workshop, “Initially, I was apprehensive as I hadn’t taught a classroom full of students for a long time. However, it was truly inspiring to work with Data Steward, [Kees den Heijer](#). We received positive feedback and I can’t wait to do it again!”

Branchett highlights the benefits of Software Carpentry workshops for researchers of all career stages. “The course teaches basic programming skills to help early career researchers get off the ground. It also offers more formal training for established researchers who may be self-taught in software development.” She adds, “Participants have an opportunity to learn new skills in a supportive environment and meet like-minded people who they can contact for help.”



Caption: Branchett instructs her first Software Carpentry workshop.

Reaching out to researchers

To reach the wider research community Branchett organised an interfaculty micro-sabbatical. This means that she spends one day per week working closely with researchers to understand how they use IT and how she can help them apply IT solutions for more innovative research. Her efforts promise campus-wide impact as she aims to spend six months working in each faculty.

Her first micro-sabbatical began in October 2018 in the [Faculty of Civil Engineering and Geosciences](#). She spent each Tuesday working with [Jerom Aerts](#), a hydrology PhD student supervised by [Nick van de Giesen](#) and Data Champion, [Rolf Hut](#), on '[eWaterCycle II](#)'. This project aims to understand water movement on a global scale to predict floods, droughts and the effect of land use on water balance. The group use a hydrological model to evaluate water flow in a real-world environment. Branchett contributed to the development of a Python [Jupyter Notebook](#) that makes data visualisation easier for the group. "I helped develop code for their model so that they can produce graphs and statistics more efficiently. Now, they can import files and generate graphs immediately without having to process data, saving both time and effort." Read more about Branchett's first micro-sabbatical project: ['Using Jupyter to study Earth'](#).

Her current micro-sabbatical in the [Faculty of Applied Sciences](#) involves installing and running Open Source software on the [Chemical Engineering](#) computer cluster. As this project ends in October, Branchett is seeking a new faculty to join. If you'd like to use her IT support in your faculty, feel free to contact her! (Details provided below).

A community for collaboration

Branchett is a [core team member](#) of the [Netherlands Research Software Engineer Community \(NL-RSE\)](#), an initiative taken by the [Netherlands eScience Center](#) and [ePLAN](#) to build a community of research software engineers in the Netherlands.

Branchett believes that fostering a collaborative community of software engineers is important for the future of science. “As our reliance on software grows, researchers possessing the technical capacity to write and contribute to software become intrinsic within the academic system. We must give these researchers more visibility and harness their expertise in order to build a sustainable scientific community.”

The second NL-RSE Conference will take place in November, inviting expertise from Dutch Universities, research institutes, companies and other organisations to facilitate networking, knowledge exchange and recognition for research software. The [call for contributions](#) to the programme is now open.

Computing Questions?

Want to do IT better? Need help choosing the right tool for the job? Want to arrange a micro-sabbatical within your research faculty? Contact [Susan: S.E.Branchett@tudelft.nl](mailto:S.E.Branchett@tudelft.nl) or visit her at the [University Corporate Office](#). You can also learn more about her interests and expertise by visiting [Twitter](#), [LinkedIn](#) and [Github](#).

How to build a community of Data Champions: Six Steps to Success.



Here, we present a ‘toolkit’ to help you build a community of Data Champions within your academic institution

Getting Started

Step 1. Preparation

Prepare to launch your Data Champions programme by discussing ideas with your research community and institutional management.

Step 2. Coordination

Coordinate a bottom-up, community-led programme to build a sustainable support network.

Step 3. Promotion

Promote your programme and make the call for volunteer participation.

Maintaining momentum

Step 4. Introduction

Introduce new recruits to your programme by giving them a warm welcome.

Step 5. Communication

Encourage collaboration and networking amongst your community by organising regular events, meetings and use of communication tools.

Step 6. Recognition

Reward and recognise the efforts of your Data Champions by publicising their achievements.

Getting Started

Step 1. Prepare to launch your programme.

Establishing a community of volunteers can be a daunting prospect. It takes time, effort and resilience to build a strong community and you may have several burning questions about how you can get started. Here's a list of 'toolkit essentials' you'll need to consider before launching your Data Champions programme...

- *A clear vision*

When you imagine your future community-based model, what do you see? It's important to outline what you expect from your Data Champions and plan how you're going to support them in their new role. Having a clearly defined outcome will help you direct your aims and objectives as you prepare to kick-start your programme.

- *A realistic goal*

Be ambitious but set realistic expectations of what you and your Data Champions can offer the programme in terms of time and resources.

- *A positive attitude*

All volunteer programmes encounter challenges. Attracting Data Champions and maintaining motivation amongst the existing cohort can prove difficult. You'll need to be versatile; be prepared to adapt your expectations and mitigate challenges as your programme develops.

- *A willing community*

Start engaging with members of your local research community. Build relationships with individuals who show passion and willingness to help you realise your ambitions. Focus groups can be a useful way to meet new people, introduce your idea, stimulate discourse and gain valuable feedback on how to start creating your community of Data Champions.

Talking to Department and Faculty managers, or data support, can give you valuable insight into the current research culture and give you inspiration to start your programme. You may also need to approach institutional management to pitch your idea, and apply for funds and resources to organise events to kick-start your programme. For example, you may need to allocate time to a member of staff to coordinate your programme. In our experience, this can range from 0.05 (TU Delft) to 0.3 FTE (University of Cambridge).

- *A desire to learn*

Do your homework! Learn from institutions that have pioneered the development of community-based models. See our [executive summary of plans](#) to launch our Data Champions programme at TU Delft and learn about '[Establishing, Developing, and Sustaining a Community of Data Champions](#)' at the University of Cambridge.

You can also read informative case studies shared by institutions that have successfully established 'Networks of Data Champions' in '[Engaging researchers with research data management: The Cookbook](#)' (Chapter 5).

Step 2. Coordinate a bottom-up, community-led programme.

At TU Delft, our Data Champion programme is centrally coordinated by the [Community Manager](#) and is intrinsically linked to the [Data Stewardship](#) programme. Similarly, the [University of Cambridge](#) Data Champion programme is centrally coordinated by members of the [Research Data Management Team](#). Both programmes are community-led meaning that researchers direct the programme and cooperate to achieve locally owned visions and goals.

Choosing to coordinate your programme, be it through central or distributed coordination, will aid its sustainability. Programmes that are entirely community-led can be less sustainable particularly if the initiative is largely driven by a handful of committed researchers. What happens if they leave your institution or can no longer commit to your programme? Centrally coordinating your programme helps to mitigate this problem since dedicated personnel are at hand to provide your Data Champions with the support and guidance they need to fulfil their role, even if it's assisting with administrative tasks, such as booking venues and arranging catering for events. Centrally coordinate your Data Champions and help them to...

- ***Build and sustain a strong support network***

Well-connected with the research community and central RDM support infrastructure, Data Champions are the glue holding the research community together, helping to create a cohesive web of support to benefit all involved. As researchers volunteering within their local community, Data Champions serve as a 'community voice', representing the various perspectives of researchers and providing feedback to improve central RDM services.

- ***Reach a wider audience***

As volunteers who share their expertise and advocate proper RDM practice across a heterogeneity of research disciplines, Data Champions have the capacity to drive a culture change towards Open Science across an entire research institution. By delivering widespread support they help to reduce the workload of the central RDM support team and raise awareness of the Data Champion programme amongst the wider research community. This proves particularly useful to recruit new Data Champions and reduce the problem of frequent volunteer turnover.

- ***Capitalise on their assets***

Your programme must support Data Champions so that they can support their research community. For the programme to be sustainable, it must be community-led. This gives Data Champions the flexibility to capitalise on their strengths, effectively engage with researchers and deliver RDM support in ways they feel most comfortable. Taking a bottom-up approach to community building, Data Champions can feel motivated to make their individual contribution to the programme and the field of Open Science.

Step 3. Promote your programme.

Once you're ready to advertise your programme you must decide how you want to recruit your best candidates. Here's some advice for how to promote your programme and make the call for participation...

- *Be explicit*

When you draft your recruitment message specify the details of your programme. Make it clear who can become a Data Champion, what the role entails and the level of commitment involved. Create a short and simple advert that communicates the need for volunteers to join your community. Don't forget to include details on how to apply (e.g. via an online [application form](#)) or where to seek further information (provide contact details!)

- *Be enthusiastic!*

Use compelling language that elicits enthusiasm in your reader. Remember to advertise the benefits of becoming a Data Champion. You're asking researchers to dedicate their valuable time and skills to your programme so make them aware of the benefits to incentivise them to participate. Take a look at the University of Cambridge's [recent call](#) for inspiration.

- *Tailor your advert*

Think about your audience. Who do you want to recruit? Take a parallel approach using broad and targeted recruitment as it can be difficult to appeal to all research disciplines using a single advert. To engage individuals with a specific expertise, share your advert across the entire institution but tailor the language to target the discipline you wish to recruit. For example, you will need a different advert for science, technology, engineering and mathematics (STEM) vs Arts, humanities and social sciences. Be prepared to make regular calls to target specific groups to achieve equal recruitment of Data Champions from all research disciplines. This not only helps to achieve inclusion and diversity but can help to maintain healthy numbers of Data Champions.

- *Make the call*

Distribute your message using flyers, posters, mailing lists, social media and via contacts within research departments. It's also a good idea to have a section on your website to promote your programme. The [TU Delft library website](#) features a selection of web pages that advertise the [role](#), [requirements](#) and [rewards](#) of a Data Champion, as well as [how to apply](#) to join the programme.

Maintaining momentum

Once you begin recruiting Data Champions, you'll need strategies in place to help you grow and maintain your network over the long-term. Explore ways to introduce new recruits, and encourage networking and collaboration to develop a vibrant and dynamic support network.

Step 4. Introduce your Data Champions.

Data Champions are your most valuable resource so it's important to give them a sense of belonging. As you introduce new Data Champions to your team, give them...

- *A warm welcome*

Respond to new recruits quickly by sending a friendly and enthusiastic welcome email. Your first contact should thank them for signing up to participate in your programme and provide essential information about the programme to acquaint them with their new role. You may choose to provide new Data Champions with a welcome pack containing the information and resources they need to get started. A welcome pack can include details of the programme's aims and objectives; training options and communication tools; plus, a list of useful online resources and key contacts. Providing an RDM glossary can also help your Data Champions improve their own data management knowledge as they take on their new role.

- *A formal induction*

In some cases, a formal induction to the programme is preferable to introduce new Data Champions to the existing network. It provides them with an opportunity to meet one another, start networking and hear first-hand accounts from established Data Champions. Recently, the University of Cambridge have introduced a pilot mentoring project as part of their induction process whereby established Data Champions volunteer to support new Data Champions. You can [read more](#) about this pilot and other exciting news.

- *A personal profile*

We create a personal profile for each new Data Champion and publish it to '[Our Data Champions](#)' page. Their profile details their area of expertise, motivations for becoming a Data Champion, contact details and photo ID which allows the wider research community to identify their local advocate for good data management and 'put a face to a name.'

Step 5. Encourage communication.

Communication is the key to creating a collaborative open research community. As part of an interactive network, Data Champions can feel empowered to recognise salient issues, find common ground for action and cooperate to implement their ideas and decisions. Here are some tips, tricks and tools to encourage communication amongst your network...

- *Organise regular meetings*

Bring your Data Champions together for regular meetings to keep your community up-to-date with the latest Open Science movements. TU Delft's meeting agendas typically consist of talks delivered by Data Champions and invited guest speakers to exchange RDM knowledge and move the programme forward.

Allow sufficient time for discussions in between talks and organise an informal networking session over refreshments to conclude your event. This always goes down a treat!

Be inclusive! Consider making your Data Champion meetings open to anyone interested in the work of your Data Champions. This can help to spread the word of proper RDM practice and drive a culture change within your institution.

- *Engage the wider community*

Create blog posts to summarise your event; share photos and video content to inform and inspire your research community. TU Delft regularly posts news updates on various Data Champion and RDM events on our [Open Working blog](#). Check out the summary of TU Delft's [Data Champions kick-off meeting](#).

- *Host an annual event*

Celebrate the success of your Data Champions each year during a one-day event to reflect on their highlights and propose new programme agendas.

- *Use communication tools*

Facilitate networking and collaboration over dedicated email lists or online communication platforms, such as [Slack](#), a software for connected teams. Give your Data Champions access to a shared environment, such as [Google Drive](#), so that they can use and contribute to your selection of programme materials, such as meeting minutes and presentations.

Step 6. Reward and recognise your Data Champions.

Recognition is critical for volunteer retention. If your Data Champions feel valued and appreciated, they'll feel better connected to the program and incentivised to participate. Make it a principal objective to commend your Data Champions on their inspiring work and express your gratitude to those that 'go the extra mile'.

- *Publicise their achievements*

There are various ways to recognise the efforts of your Data Champions and they don't have to be resource-intensive.. For some, simply joining the community of Data Champions and having their personal profile published on the community website is reward enough.

When a Data Champion joins your programme it may also be a good idea to notify their Head of Department or Faculty Dean. This can help to raise awareness of your programme more generally and introduce your Data Champion as a new source of local RDM support.

You may also choose to reward the efforts of your Data Champions by giving them more visibility which can raise their professional profile. A current project at TU Delft publicises the achievements of our Data Champions by interviewing them to learn more about their research projects; how they effectively engage with researchers; their motivations for becoming Data Champions; and, their future goals and aspirations. Following each interview, each Data Champion case study is written and published as an article on TU Delft's Open Working blog under a dedicated tab on the Homepage titled '[Data Champions](#)'. The articles are promoted on social media to put Data Champions in the spotlight and increase impact beyond their immediate circle.

Regular one-to-one engagement with Data Champions also allows you to gain their personal perspective on the programme and measure the impact of your programme on driving a culture change towards Open Science.

- *Provide travel opportunities*

If funds are available, give your Data Champions an opportunity to apply for travel grants to attend international conferences. TU Delft Data Champions are eligible to apply to a dedicated 'Data Champions Travel and Training Fund' to attend workshops and conferences related to data stewardship. During the first year of our programme we have [funded two Data Champions](#) to attend international workshops with the total cost below €1,500.

Aside sharing their RDM expertise at these events, Data Champions can engage with representatives of similar programmes, learn new skills and expertise, and disseminate knowledge gained back to your Data Champions programme.

- *Nominate exemplary candidates for awards*

Recognise excellence and acknowledge those that go above and beyond for their research community by nominating them for internal or external awards, such as the [WDS Data Stewardship Award](#).

Encourage their participation in other international initiatives, for example, the [SPARC Europe's Open Data Champions](#), and support their applications by writing professional references and letters of support. These extracurricular activities promise to boost their CV!

Based on the premise that recognition is a vital aspect of sustaining a community of Data Champions, we continue to explore novel ways to reward our Data Champions. A challenge is understanding how to reward Data Champions at different career levels. As you establish your own volunteer community, we welcome any suggestions you might have to inspire and help us along the way! We wish you success in your future endeavour.

A message from the author:

I would like to express my grateful thanks to TU Delft Data Champion Community Manager, [Yasemin Türkyilmaz-van der Velden](#), and Data Stewardship Coordinator, [Marta Teperek](#), for their support and advice during the development of this toolkit.

In addition, thanks to [Lauren Cadwallader](#) and [Sacha Jones](#) from the University of Cambridge, [Francesco Varrato](#) from École polytechnique fédérale de Lausanne, and [Hardy Schwamm](#) from NUI Galway for your constructive feedback and for inspiring some of the content.

