Introducing Mendeley Data: Supporting the Lifecycle of Research Data Management in Support of Open Science
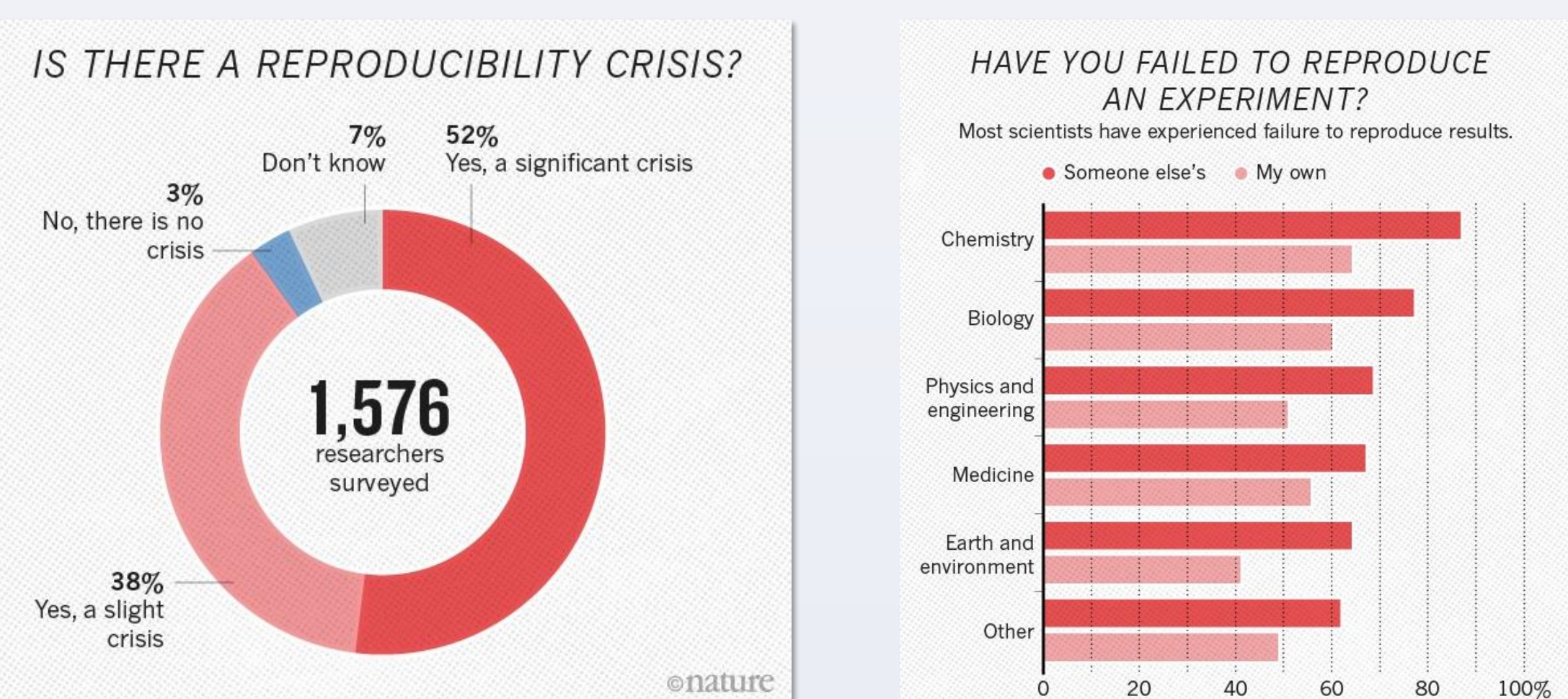
Holly J. Falk-Krzesinski, Alberto Zigoni, Helena Cousijn, Anita de Waard, Peter Cotroneo, Elena Zudilova-Seinstra, Jean P. Shipman, Rebecca Brown, Wouter Haak

Research Data Management team, Elsevier

Mendeley Data

ELSEVIER

Data.Mendeley.com

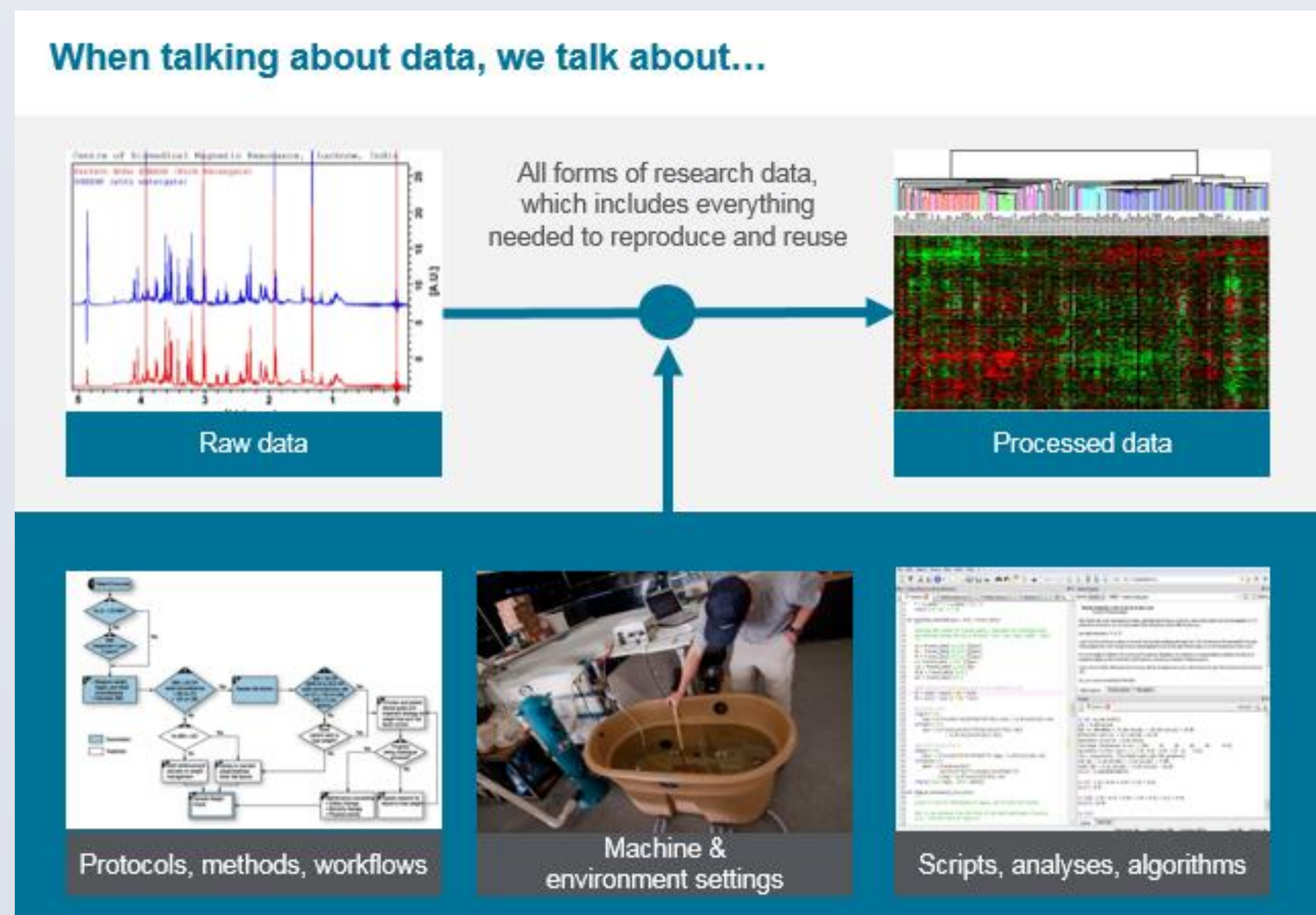## Supporting researchers & institutions increase reproducibility

Research data sharing is a cornerstone for success in modern science. All the information underpinning research articles has value to other researchers: raw and processed data, protocols and methods, machine and environment settings, and scripts and algorithms. This has driven international funding bodies to revise their policies to encourage, or even mandate, institutions and their researchers to make research data available.

IS THERE A REPRODUCIBILITY CRISIS?

7% Don't know | 52% Yes, a significant crisis
3% No, there is no crisis
1,576 researchers surveyed
38% Yes, a slight crisis

HAVE YOU FAILED TO REPRODUCE AN EXPERIMENT?
Most scientists have experienced failure to reproduce results.
Someone else's • My own
Chemistry, Biology, Physics and engineering, Medicine, Earth and environment, Other

Source: https://www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970

Recognition of the value and necessity of data sharing has driven the development of secure and stable online solutions to make this information permanently available for interpretation and reuse.

From research data plan to active collaboration

Of course, simply depositing data in folders on a shared server isn't enough to ensure that researchers can derive the most value from them. They must be accurate, organized, searchable and accessible from multiple locations, but also secure and held in a stable environment. When data are easy to find and share, that facilitates collaboration and ensures knowledge is not lost.

When talking about data, we talk about...

All forms of research data, which includes everything needed to reproduce and reuse

Raw data | Processed data
Protocols, methods, workflows | Machine & environment settings | Scripts, analyses, algorithms
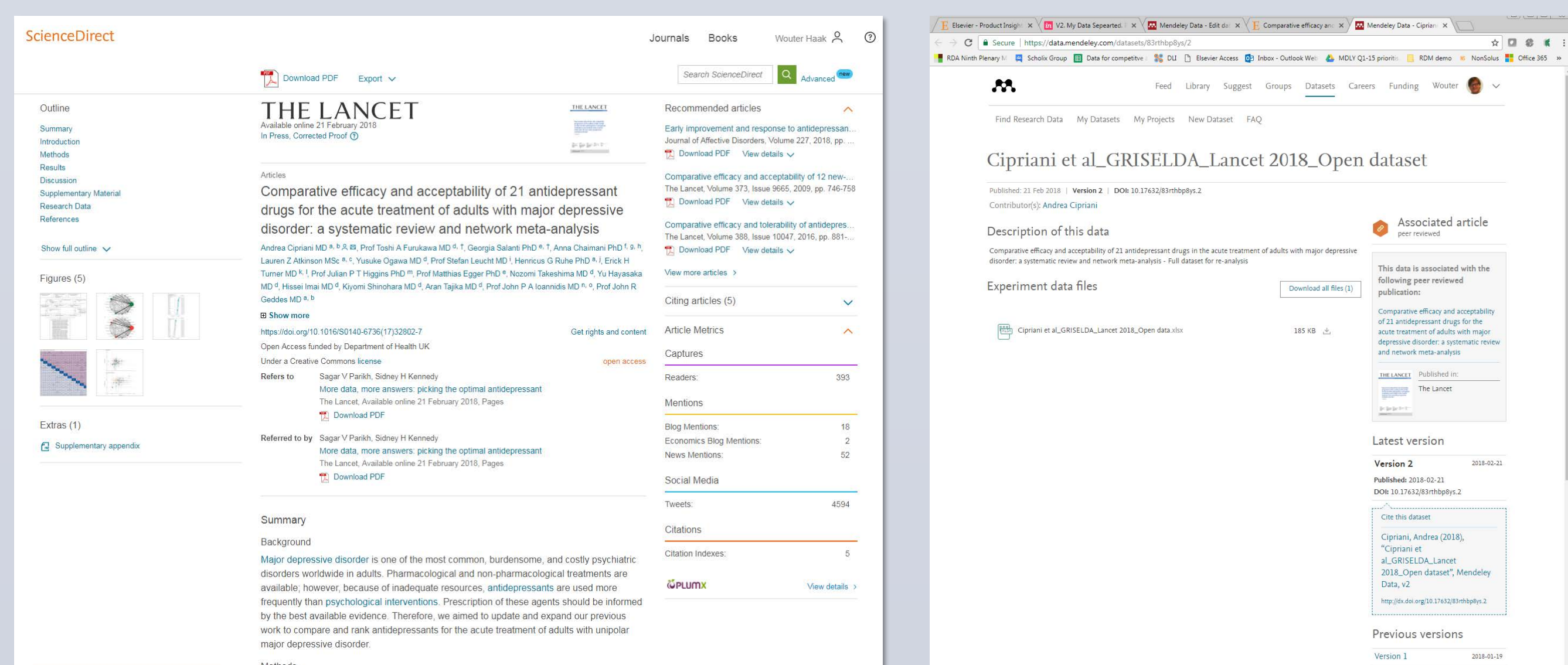
As data management and sharing requirements have evolved, we've been refining and improving our data management solutions. Specifically, we've examined how we can ensure that we've provided the right tools to:
- Capture and index data in the lab so that nothing is lost
- Organize and annotate data so that it's ready for sharing and publication
- Securely store and share data in the right formats for collaboration
- Facilitate the searching of data so that it's always discoverable
- Oversee the management of data so that workflows remain compliant

Over 8 million datasets are currently discoverable on Mendeley Data

We all want to see more relevant data be shared & used, and its our mission to work with institutions to make it easier and more rewarding for researchers to do so. Our free search tool enables easier and more accurate searching thanks to deep indexing of data (not just metadata) sourced from over 30 repositories & growing- see just one example below.
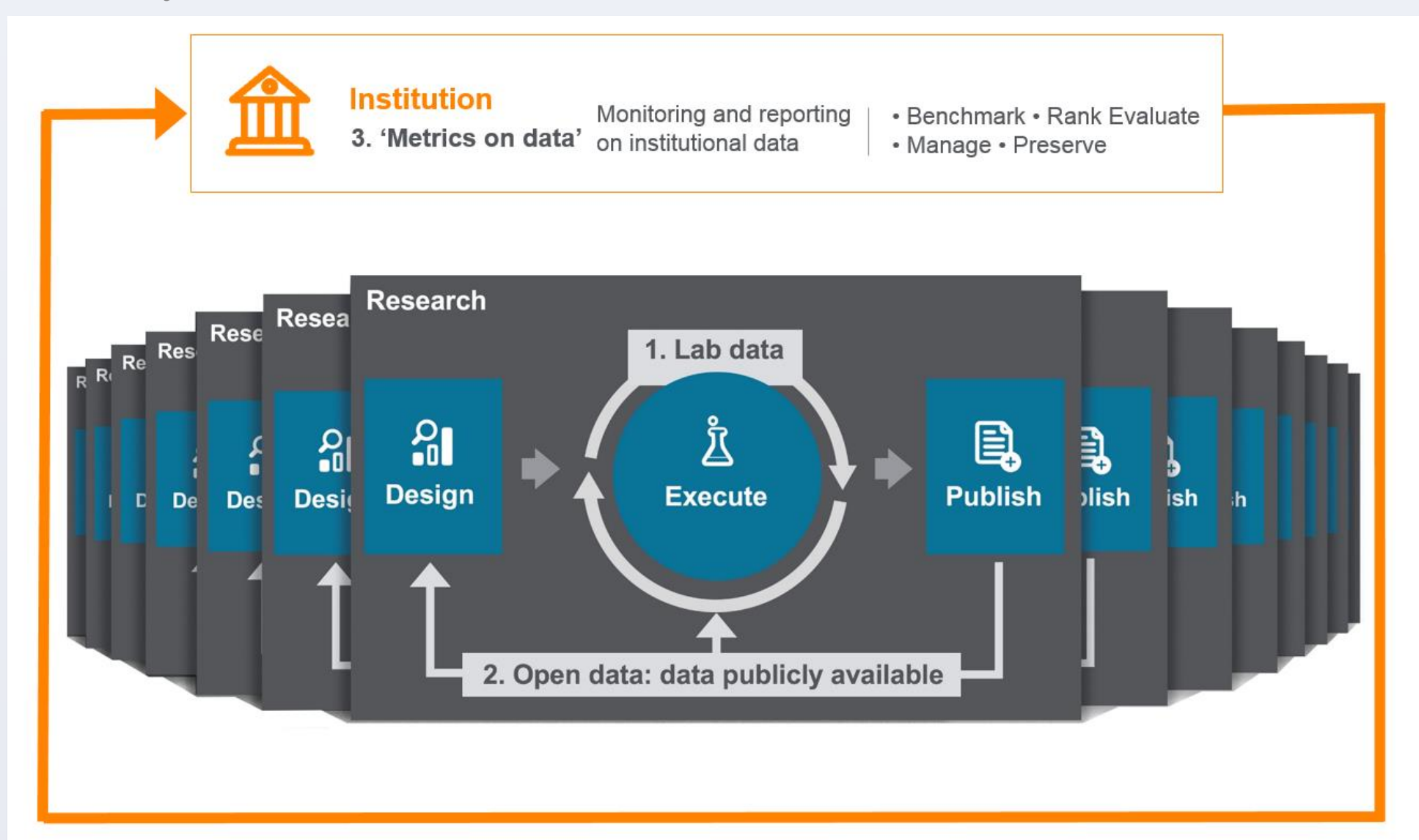

Example: Lancet article & dataset on the comparative efficacy & acceptability of antidepressant drugs from Cipriani et all, Accessed 21 Feb 2018

This is just a start, as we know managing data at an institutional level means that you need to have tools that are already used by researchers; we have been working with community and development partners on expanding to meet the institutional need.
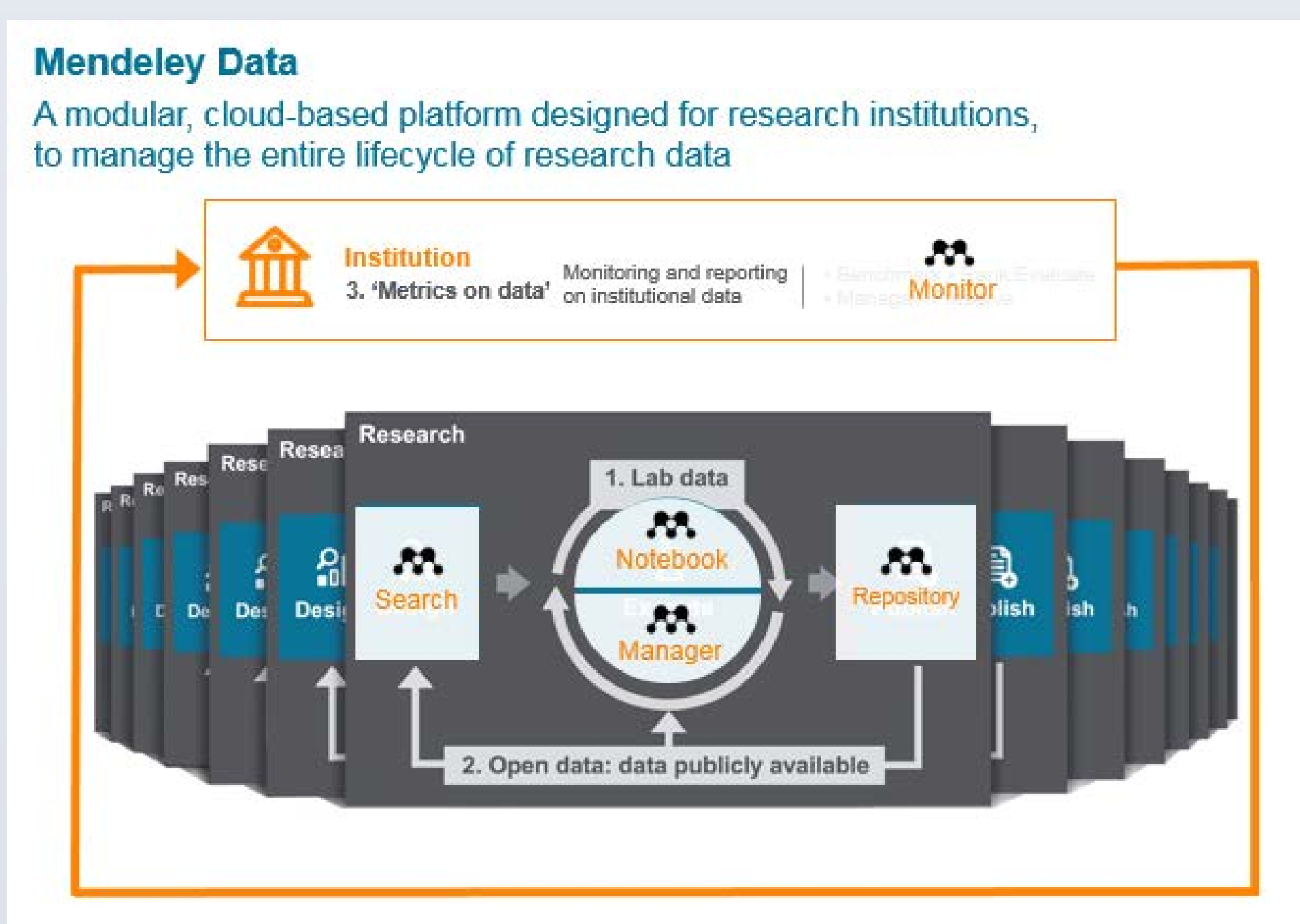

Introducing Mendeley Data

Many will recognize this image– it's the *primary research lifecycle*. Researchers start in the design phase by finding what project to work on next & getting the funding for it. They move into execution mode, carrying out the research itself and then finally publishing and sharing results with their networks. The research lifecycle depends on two *data lifecycles*: one for all active data within the lab and one for all publicly available data. From an institutional point of view, there is a third lifecycle: that of the *metrics on data*.

Mendeley Data covers key aspects of all 3 lifecycles, providing you with tools that support in-lab data recording; sharing and management; external data publication and other collaboration; and institutional monitoring of data usage. All the tools are available through a highly user-friendly interface.

Institution 3. 'Metrics on data' — Monitoring and reporting on institutional data • Benchmark • Rank Evaluate • Manage • Preserve

Research — 1. Lab data — Design — Execute — Publish

2. Open data: data publicly available

Mendeley Data is designed to facilitate the comprehensive utilization of data. Consisting of five modules, this open, cloud-based platform helps research institutions to manage the entire lifecycle of research data and enables researchers to safely access and share information wherever they are. These modules help institutions and researchers to collaborate, more easily tap into the broader world of shared data, increase research exposure and support compliance.

Mendeley Data
A modular, cloud-based platform designed for research institutions, to manage the entire lifecycle of research data

Institution 3. 'Metrics on data' — Monitoring and reporting on institutional data — Monitor

Research — 1. Lab data — Design — Search — Notebook Manager — Repository — Publish

2. Open data: data publicly available

What are the benefits of Mendeley Data for users?
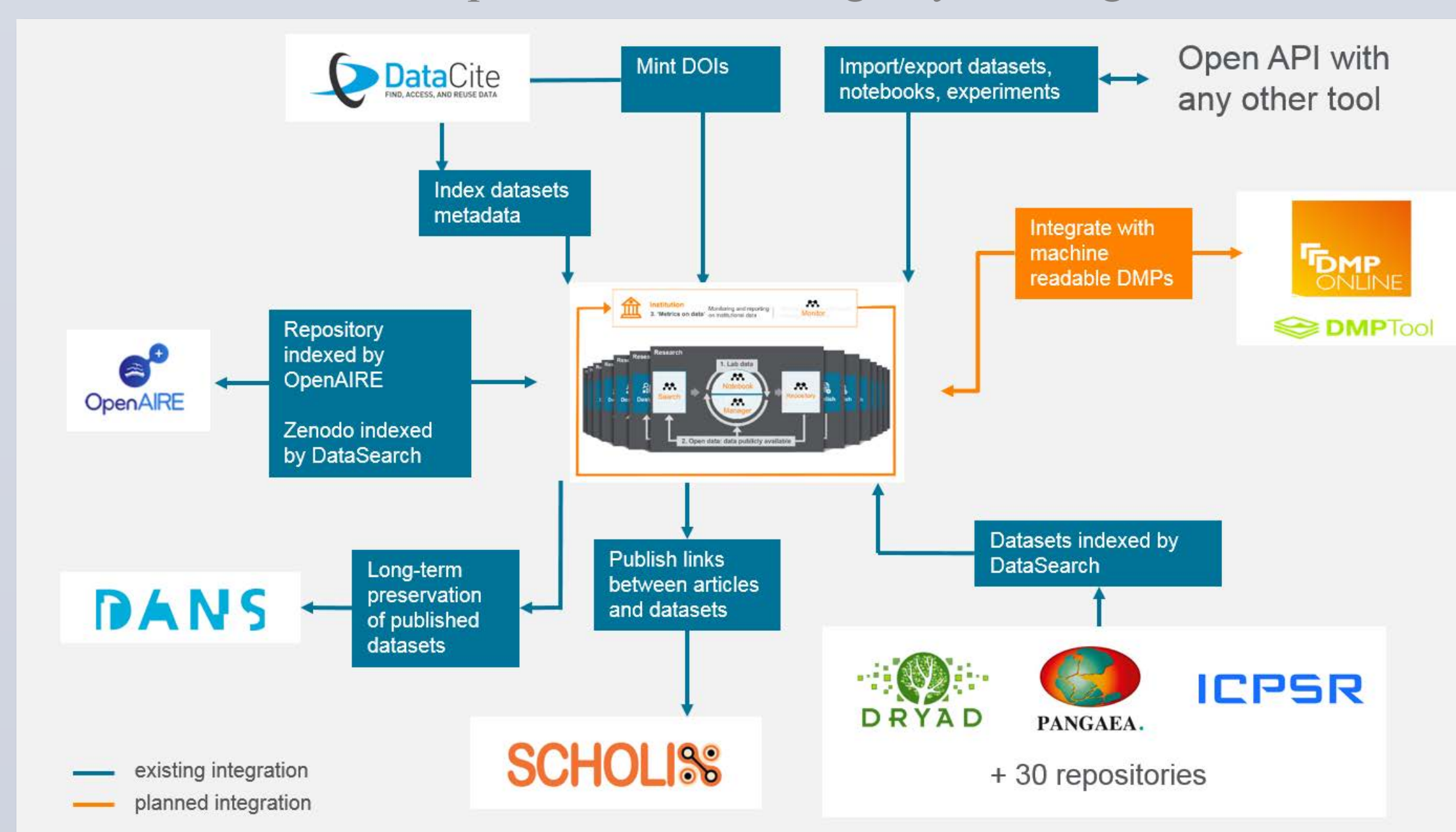
Benefits for institutions:
- **Keep track** of your data inside and outside your institution
- **Save time** and resources with easier data management
- **Showcase** institutional research outputs
- **Improve** collaborations within and between institutions

Benefits for researchers:
- **Eliminate re-work** thanks to better access to protocols, methods, etc.
- **Easily comply** with funders' mandates
- **Improve impact** through increased re-use of data
- **Delve into the details** of interesting research

An open platform: Mendeley Data integrates with the ecosystem

Mendeley Data integrates through open APIs with the global ecosystem for research data management. We link to DANS for long-term data set preservation, DataCite for DOIs and indexed metadata to support data publication, Scholix for links between published articles and datasets, and over 30 data repositories, including Dryad, Pangaea and ICPSR.

DataCite — Mint DOIs | Import/export datasets, notebooks, experiments — Open API with any other tool
Index datasets metadata
Integrate with machine readable DMPs — DMP ONLINE — DMPTool
Repository indexed by OpenAIRE | Zenodo indexed by DataSearch
OpenAIRE
DANS — Long-term preservation of published datasets
Publish links between articles and datasets
Datasets indexed by DataSearch
SCHOLIX
DRYAD | PANGAEA | ICPSR
+ 30 repositories
— existing integration
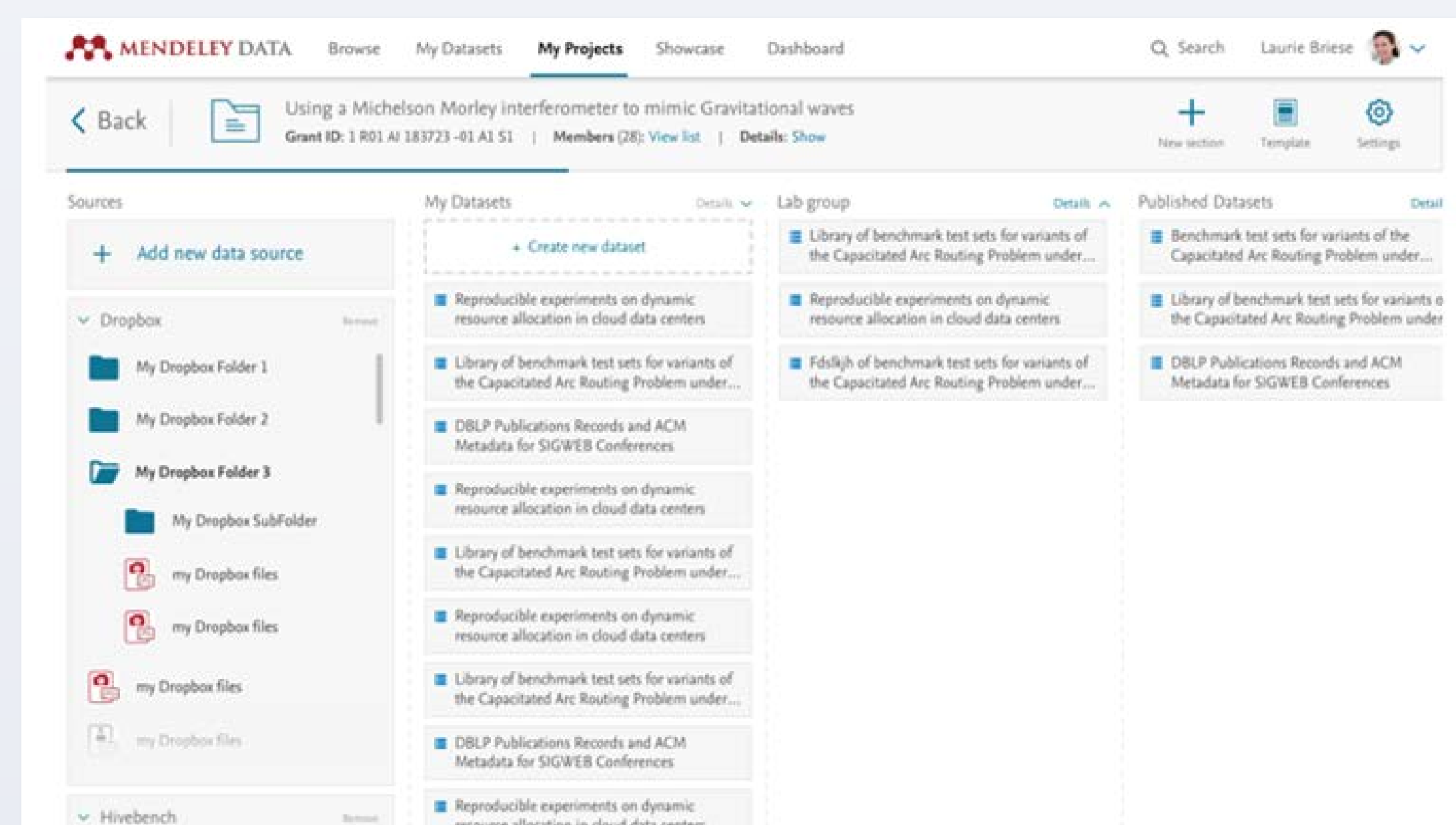— planned integration

We won't stop here: we already have planned future integration with machine-readable data management plans and other electronic lab notebook software, just to name a few developments.

Manage & Collaborate on Active Data: Mendeley Data Manager
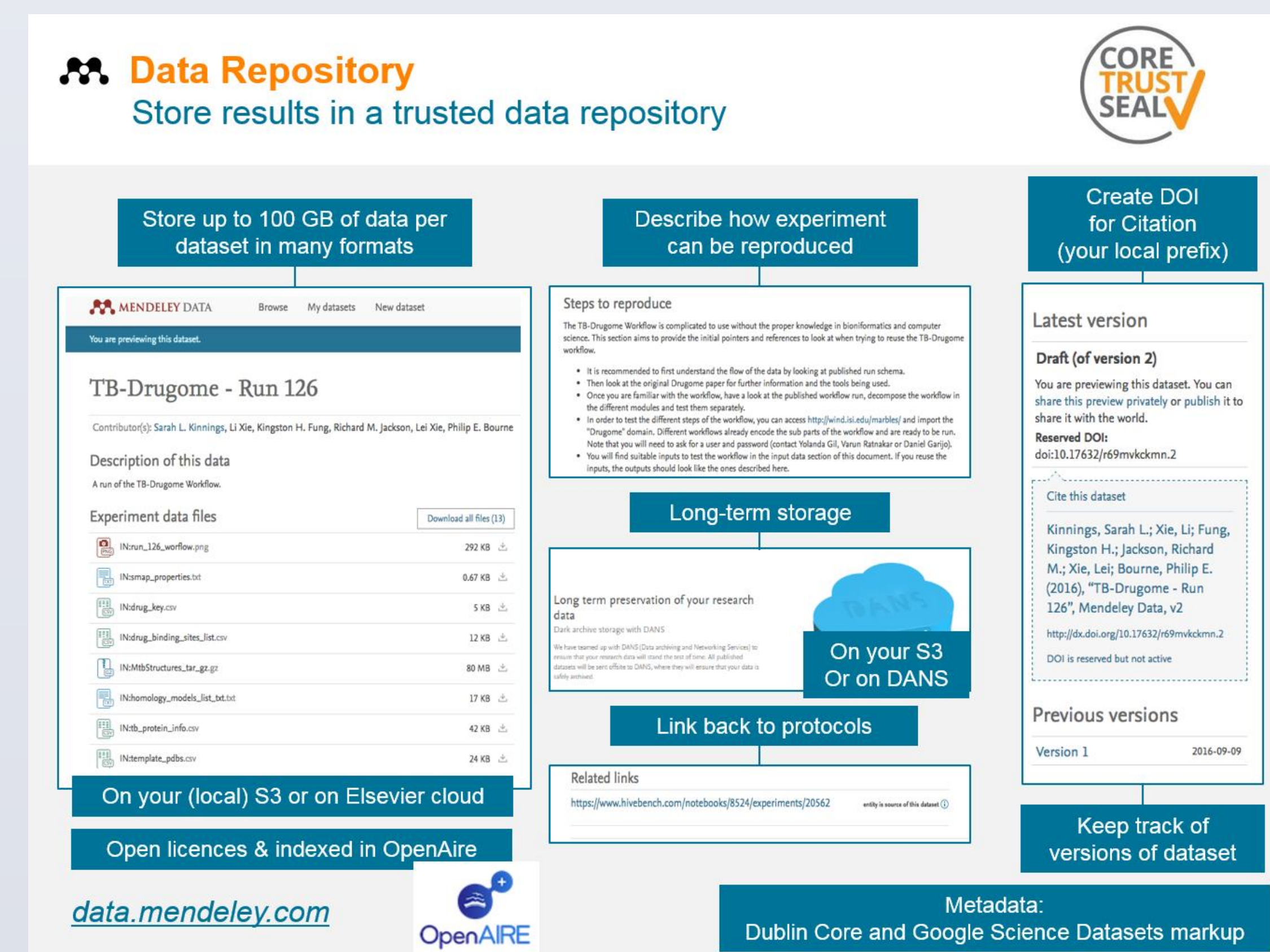
Go from raw files to published datasets

Mendeley Data Manager is an active research data collaboration and workflow tool that enables individual researchers and groups to organize, annotate and share data all in one place.

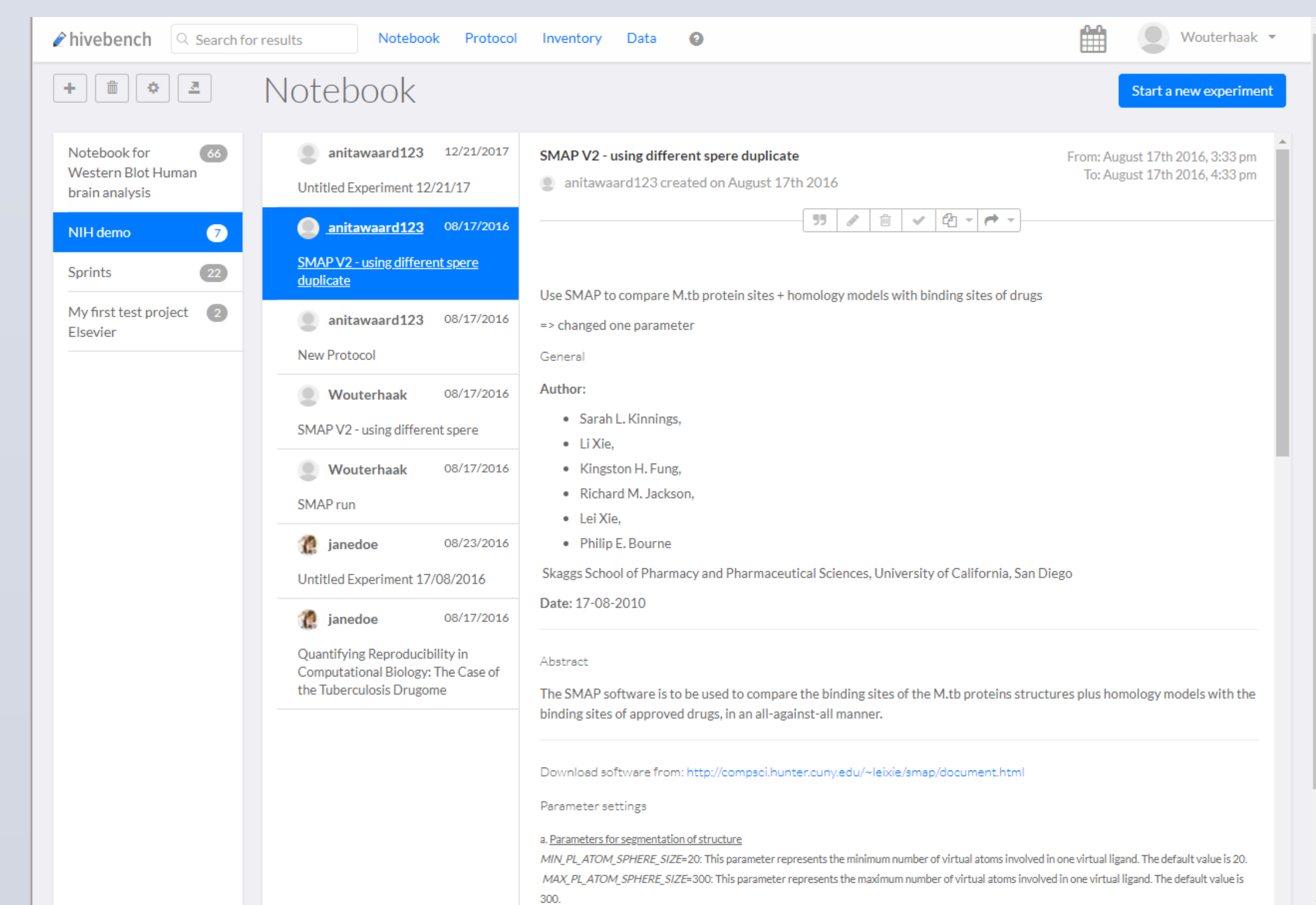### Store & Showcase Published Data: Mendeley Data Repository

Store datasets in a secure & trusted repository

Mendeley Data Repository is the institutional module that enables all researchers within a given institute to post, share, archive, showcase and discover research data.

Data Repository — Store results in a trusted data repository — CORE TRUST SEAL

Store up to 100 GB of data per dataset in many formats | Describe how experiment can be reproduced | Create DOI for Citation (your local prefix)

TB-Drugome - Run 126
Long-term storage
On your S3 Or on DANS
Link back to protocols
On your (local) S3 or on Elsevier cloud
Open licences & indexed in OpenAIRE
data.mendeley.com — OpenAIRE
Keep track of versions of dataset
Metadata: Dublin Core and Google Science Datasets markup

Capture Data & Results in the Lab: Mendeley Data Notebook

Mendeley Data Notebook is a cutting-edge electronic lab notebook (ELN) solution that supports data capture and reproducibility, improves researcher efficiency, and enables one-button reporting to all levels of the organization.

Want to learn more?

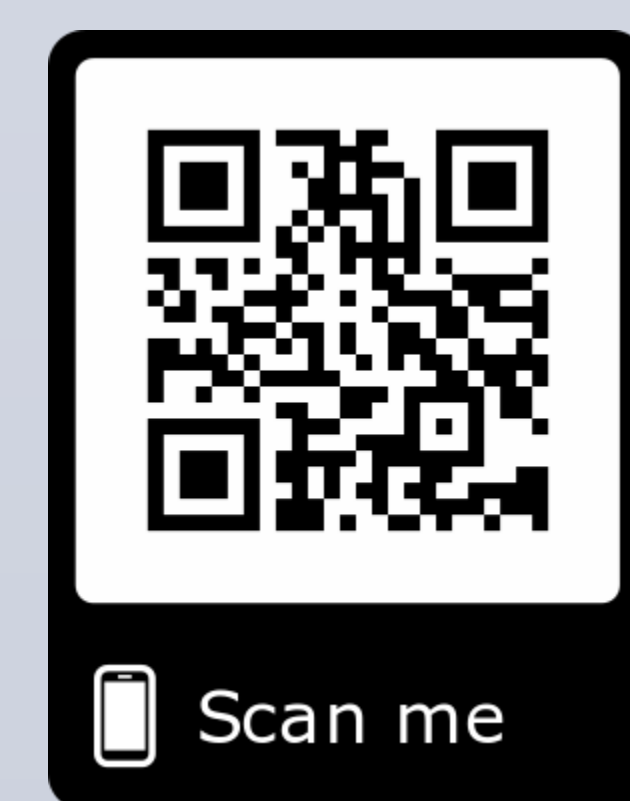Visit Mendeley Data now and get started! Or want to request a consultation?

www.elsevier.com/solutions/mendeley-data-platform

Contact us on Twitter or LinkedIn
@Mendeley_com
@Mendeley_Data
LinkedIn: Research Intelligence

Scan me

Mendeley® Data

Introducing a modular, cloud-based platform that helps you manage the entire lifecycle of research data.

Ready access to research data is a cornerstone for success in modern science. All the information underpinning research articles has value to researchers: raw & processed data, protocols & methods, machine & environment settings, and scripts & algorithms. International funding bodies have revised their policies to suggest or even mandate that institutions and their researchers make research data available.

Mendeley Data is designed to help you utilize data to its fullest potential. It is a secure, stable, cloud-based platform consisting of five modules to help you manage the entire lifecycle of research data.

Mendeley Data plays a role at the individual user and institutional levels

Researchers can:
- Avoid re-work with better access to protocols, etc.
- Easily comply with funders' mandates
- Improve impact via increased data re-use
- Delve into the details of research

Institutions can:
- Keep track of data
- Save time and resources for data management
- Showcase institutional research output
- Improve collaborations within & between groups

The modules of Mendeley Data

Data Manager
Go from raw files to published datasets

Designed to support internal and external collaboration, Data Manager includes tools for organizing, annotating and sharing data.

Data Repository
Keep datasets in a trusted repository

Data Repository manages the secure storage, sharing and dissemination of data with over 100 GB of storage per dataset in multiple formats.

Data Notebook
Capture data and results in the lab

Data Notebook is a modern and efficient electronic lab notebook (ELN) that supports data capture and indexing and even allows one-button reporting.

Data Search
Retrieve active data, discover public data

Data Search enables easier and more accurate searching thanks to deep indexing of data (not just metadata) sourced from over 20 repositories.

Data Monitor
Encourage and monitor compliance

Data Monitor makes it easier to improve the adoption of data sharing by providing publisher-independent workflow tools to help librarians. That ensures better compliance with new mandates and funder regulations.

These modules help institutions and researchers to collaborate, more easily tap into the broader world of shared data, and increase research exposure and compliance.

How Mendeley Data works

The research lifecycle depends on two data lifecycles: one for all active data within the lab and one for all publicly available data. From an institutional point of view, there is a third lifecycle: that of the metrics on data.

Mendeley Data covers key aspects of all 3 lifecycles, providing you with tools that support in-lab data recording, sharing and management; external data publication and other collaboration; and institutional monitoring of data usage. All the tools are available through a highly user-friendly interface (Figure 1).
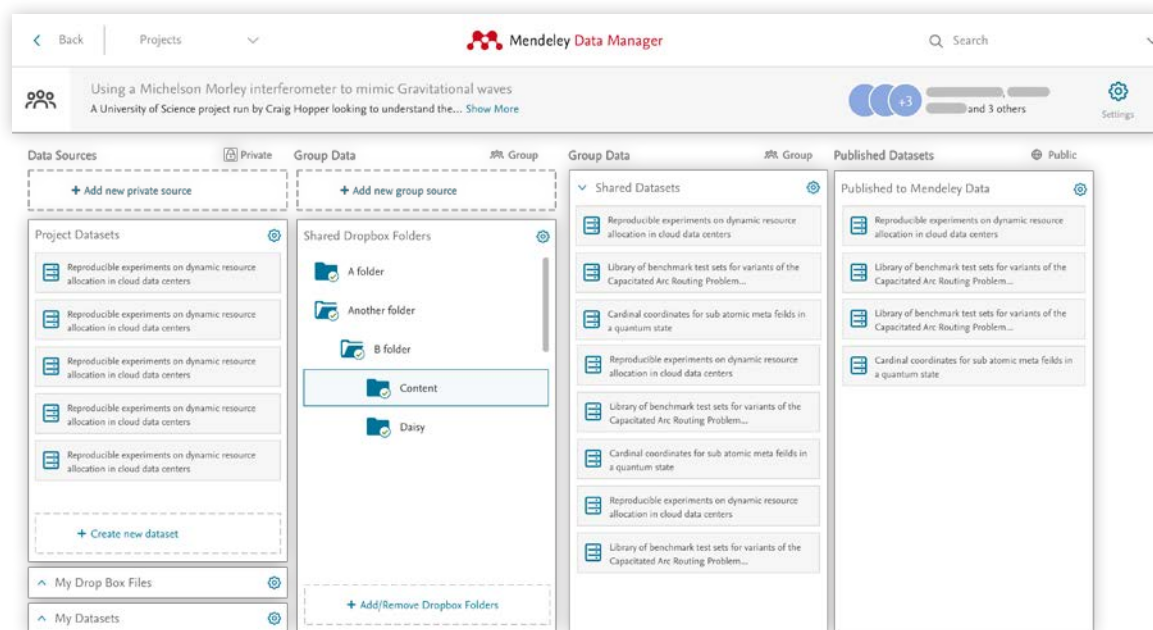

Figure 1. The Mendeley Data user interface

Mendeley Data integrates through open application programming interfaces with the global ecosystem for research data management, with links to:

- DANS for long-term data set preservation
- DataCite for DOIs and indexed metadata to help with data publication
- Scholix for support of links between published articles and datasets
- Over 20 data repositories, including Dryad, Pangaea and ICPSR

Future developments include integration with machine-readable DMPs and extension of compatibility with other ELN software.

Mendeley Data is also easy to integrate with other Elsevier research solutions, creating a stronger institutional informatics ecosystem.

- Datasets can be linked to publications in EVISE™ and ScienceDirect®
- Notifications on new articles from Scopus® help with monitoring data sharing compliance
- Datasets, projects, grants, equipment and more can be synced using Pure
- Users of the popular Mendeley platform have easy access to Mendeley Data

Future developments include Data Notebook integration with Reaxys®, production and viewing of data metrics through SciVal® and Plum Analytics, and dataset showcasing via Digital Commons®.

Put your research data online today with Mendeley Data at **data.mendeley.com**.