

The next evolution of UQ's Research Data Manager

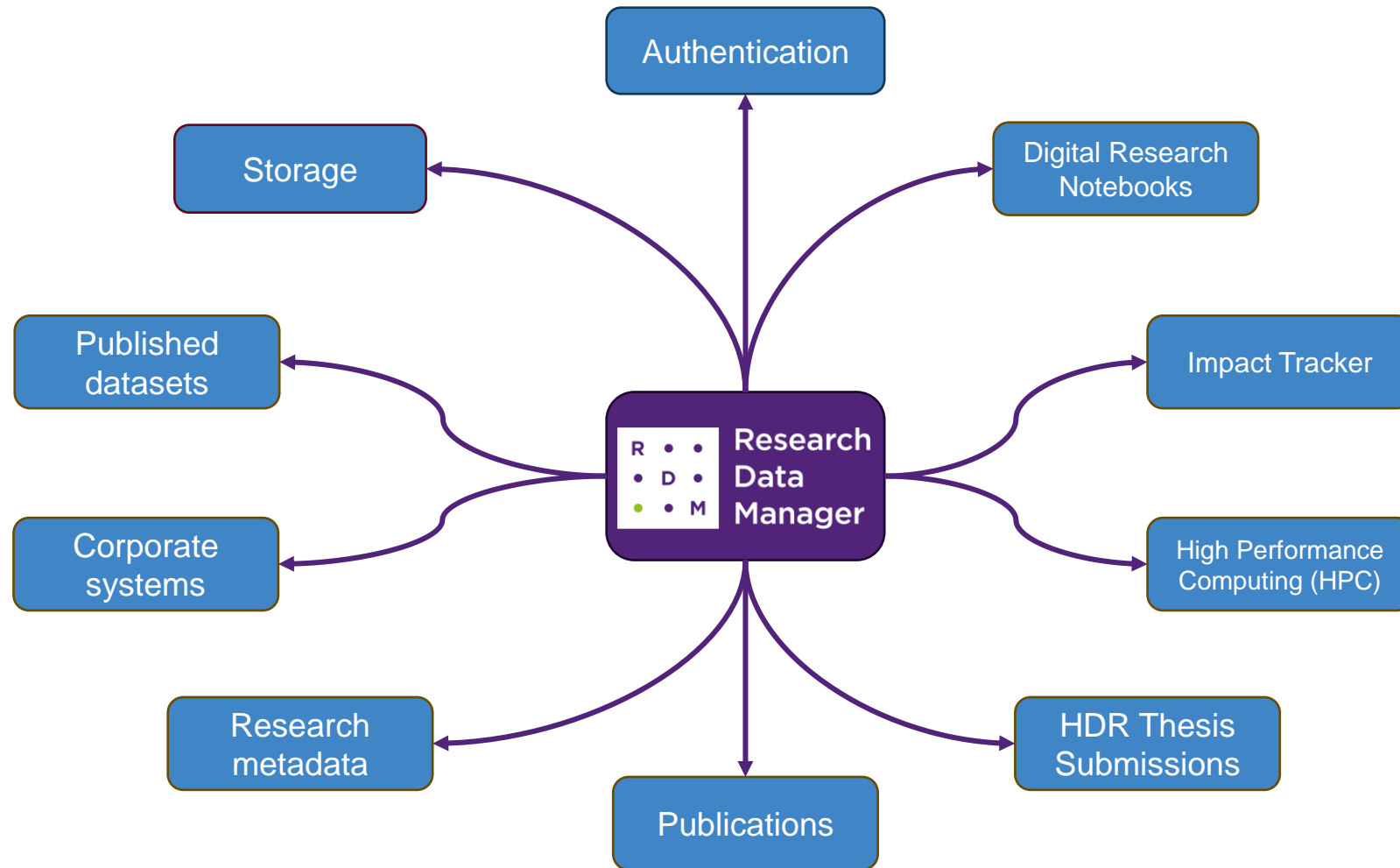
By: Jan Wisgerhof

Recap

- **UQ Research Data Manager (UQRDM) is one of the most comprehensive solutions world-wide**
- **Through our researcher-first approach, we managed coverage of ~95% of researchers at UQ**
- **Integration with third-party services is key**
- **Integration into the research project lifecycle is key**

- **Collection of metadata is almost “accidental”. Researchers are providing us with metadata without realizing it**
- **Metadata in a DMP is influenced by many factors, such as ethics approvals, grants, publications, etc**

Integration overview



**What does UQ
receive in return
for the service?**



Cyber security and compliance

- **An estimated 95%+ of all current research data at UQ is stored within UQRDM records**
- **By storing data through UQRDM:**
 - Researchers meet data storage requirements from funders
 - UQ maintains oversight and control over this data in case of “unfortunate events”
 - UQ can ensure proper cyber security procedures are applied to both storage and access to storage
 - UQ can act quickly in cases of misconduct, data breaches, etc
- **Case study:**
 - **Thursday - 9:49am: Received enquiry about research data related to a misconduct case**
 - **Thursday - 9:59am: Sent report of research data that individual has access to**
 - **Thursday - 2:45pm: Received request to remove access to research data**
 - **Thursday - 3:01pm: Access removed from all research data for individual**
 - **Friday – investigation started by storage providers into possible unauthorized removal of data**

Statistics

Field of research codes

FoR code	FoR Name	# Records	# Storage allocations	# Gigabytes	# Files
1109	Neurosciences (2008)	667	667	2,267,246	65,881,680
3209	Neurosciences (2020)	431	434	1,993,431	45,202,666
6	Biological sciences (2008)	239	239	834,009	25,881,808
601	Biochemistry and Cell Biology (2008)	299	299	830,672	27,660,277
60408	Genomics (2008)	185	185	791,116	20,025,261
3101	Biochemistry and Cell Biology (2020)	324	325	751,778	20,079,124
3102	Bioinformatics and computational biology (2020)	262	265	705,616	30,628,491
60102	Bioinformatics (2008)	226	226	695,178	21,128,358
3105	Genetics (2020)	225	231	628,340	24,584,355
10202	Biological mathematics (2008)	26	26	628,125	3,122,085
310706	Virology (2020)	112	114	599,507	15,202,138
604	Genetics (2008)	248	248	570,681	14,775,960
310204	Genomics and transcriptomics (2020)	135	139	524,827	16,964,089
310509	Genomics (2020)	169	170	492,036	19,251,611
110999	Neurosciences (not otherwise classified) (2008)	65	65	476,228	12,023,447
3204	Immunology (2020)	285	292	465,405	31,146,099
310207	Statistical and quantitative genetics (2020)	112	114	411,099	14,468,963
310112	Structural biology (2020)	134	139	394,782	9,215,624
60412	Quantative Genetics (2008)	96	97	392,003	16,779,380

Statistics

Organisational units

Name	Type	Number of active users	Number of active records	Storage used (acti
Queensland Brain Institute	Institute	557	1363	5663.94 TB
Institute for Molecular Bioscience	Institute	1015	1417	4371.74 TB
School of Chemistry and Molecular Biosciences	School	857	1122	2484.77 TB
Frazer Institute	Faculty Centre	368	734	1964.83 TB
Queensland Alliance for Agriculture and Food Innovation	Institute	375	479	876.97 TB
School of the Environment	School	1236	407	844.76 TB
Centre for Microscopy and Microanalysis	University Centre	98	297	700.14 TB
Centre for Advanced Imaging	University Centre	42	327	683.18 TB
School of Biomedical Sciences	School	786	538	624.76 TB
Australian Institute for Bioengineering and Nanotechnology	Institute	930	1187	561.89 TB

Statistics

Collaboration within UQ – Queensland Brain Institute

Organisational unit	Unique collaborators	Total collaborators
Queensland Brain Institute	409	2895
School of Biomedical Sciences	42	311
Clem Jones Centre for Ageing Dementia Research	21	185
Australian Institute for Bioengineering and Nanotechnology	43	150
Asia-Pacific Centre for Neuromodulation	5	82
Faculty of Science	53	81
School of Psychology	30	65
Office of the Vice-Chancellor	21	63
Faculty of Medicine	35	56
School of the Environment	9	52
Mater Research Institute-UQ	13	47
School of Electrical Engineering and Computer Science	20	42

Statistics

Thesis submission statistics

We make real business decisions and process changes based on data.

Review action per average thesis

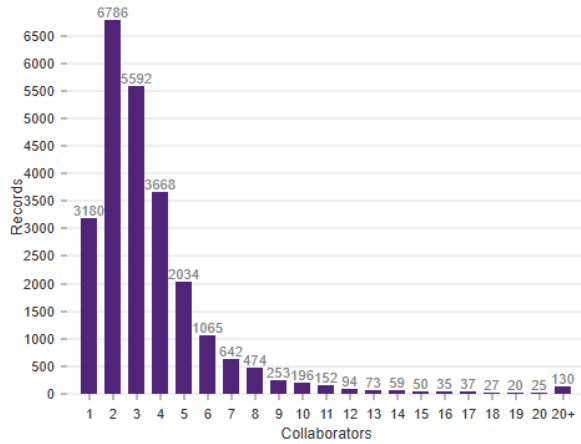
Average number of actions per submitted thesis

Action requested	Occurrence per submission (%)	Occurrence per submission (% of theses in <100 days)
Candidate withdrew	12.02 %	15.00 %
PA requested changes	18.69 %	20.00 %
GS requested changes	64.63 %	68.13 %
Revise and resubmit	0.30 %	3.13 %

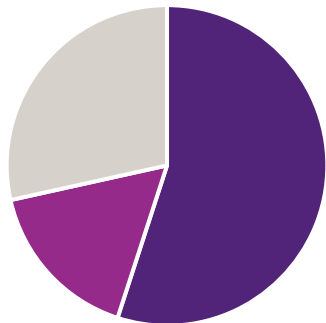
Statistics

And so much more...

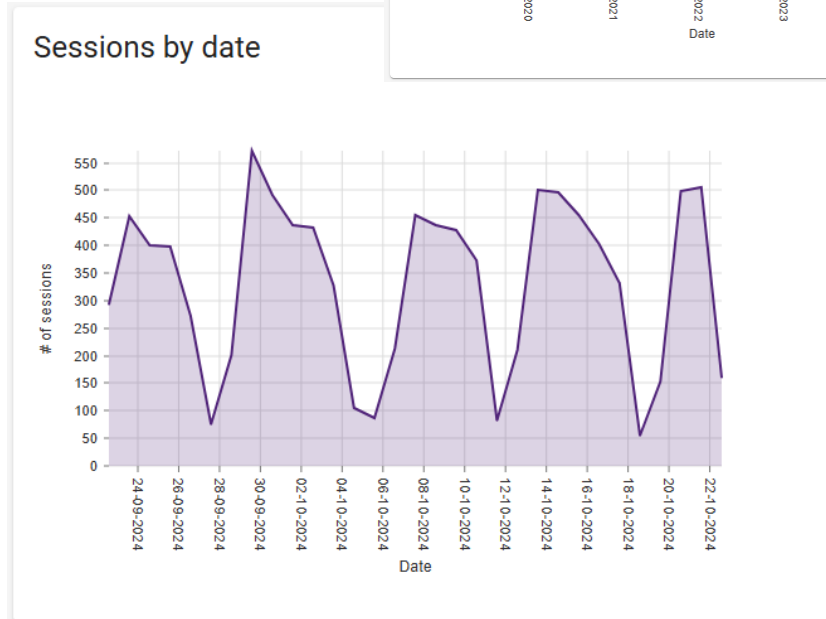
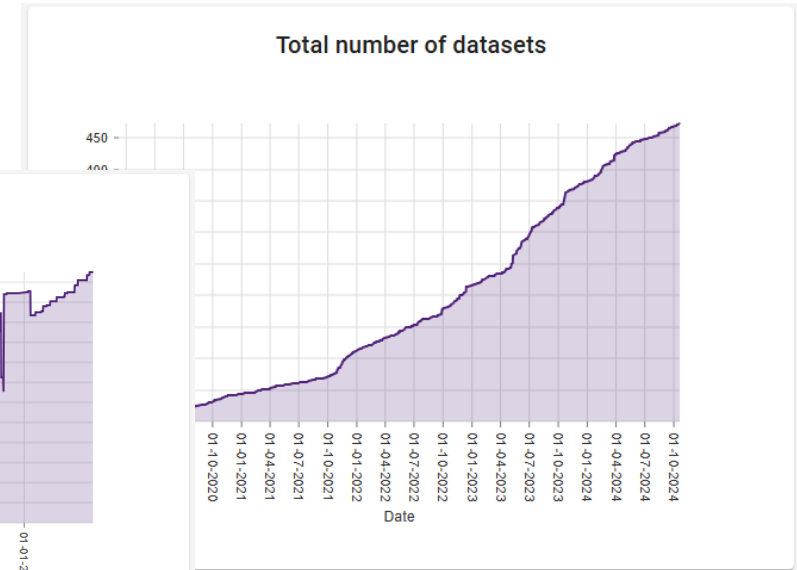
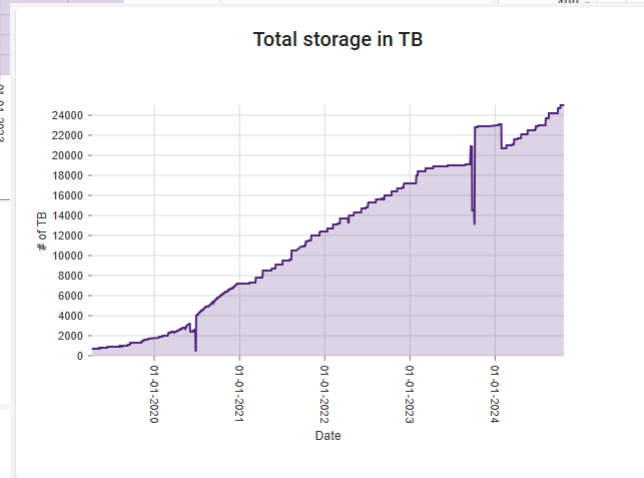
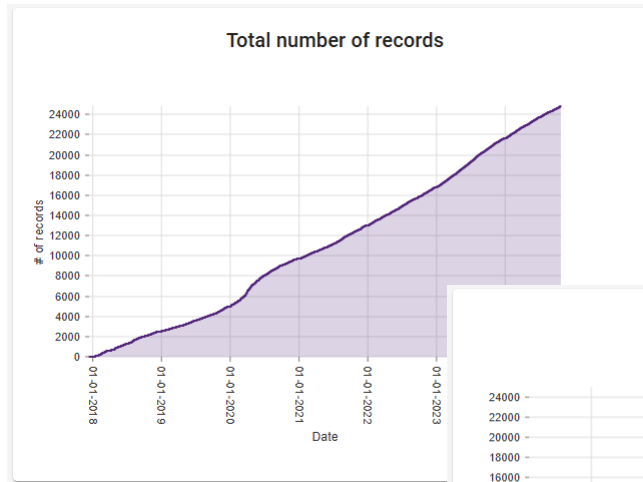
Number of collaborators on a record



Allocations by type



■ General research data ■ Human data ■ HPC enabled data



Number of users per domain

Domain	Number of users
gmail.com	8795
health.qld.gov.au	3241
hotmail.com	2556
uqconnect.edu.au	1598
mater.org.au	792
outlook.com	703
eq.edu.au	667

Provenance

- **Live, accurate provenance over who has access to research data at UQ**
- **Provenance is driven not by manual entry, but by researchers providing access to storage**
- **Provenance is visible to researchers and administrators**

Event	User	Date
Collaborator "Andrew Kneynsberg" added to vendor for service "saturday-A12002" (STORAGE) with permission: "Read / Write"	System	27-Aug-2024 11:30:14
Collaborator "Andrew Kneynsberg" requested from vendor for service "saturday-A12002 (STORAGE) with permission "Read / Write". Pending re...	System	27-Aug-2024 11:13:31
Permissions of "Andrew Kneynsberg" were changed from "No access" to "Read / Write" on service "saturday-A12002" (STORAGE)	Mr Jan Wis...	27-Aug-2024 11:12:58
Record service "Saturday Go Live" (IMPACT_TRACKER) was requested from the vendor	System	22-Jul-2024 14:40:44
Collaborator "Mr Jan Wisgerhof" added to vendor for service "Saturday Go Live" (IMPACT_TRACKER) with permission: "Chief Investigator"	System	22-Jul-2024 14:40:44

What's next?

Meeting legislative requirements

- As research data continues to come more into the spotlight, universities will increasingly be asked to report on their data holdings, their characteristics and sensitivities
- Large likelihood that UQRDM will be mandated from 2025 onwards for all funded research at UQ
- UQRDM will require researchers to fill in more metadata when requesting storage
- But we are going to do this whilst delivering new functionality for researchers through a “DMP wizard” that also meets various funder requirements, etc



UQRDM 2026+

UQRDM program of work currently being developed

- **Head of school dashboards**
- **DMP wizard**
- **HDR Enhancements (progress reviews, etc)**
- **Retention, disposal and archival workflows**
- **Dataset publication workflow enhancements**
- **More service integrations:**
 - REDCap
 - eLab Notebooks & inventory
 - Secure eResearch Platforms
 - Further HPC integration



What about 2025?

- **UQRDM integration is far ahead of most University's data management systems, but we can do better...**
- **Not all "integration" is made equal, opt-in versus automatic**
- **Thesis submissions are a great example. Because HDR students need to submit their thesis in UQRDM:**
 - We know their PA, research subject, FoRs, etc
 - We know where the research data for the thesis is stored
 - We know which grants & ethics applications are relevant to the thesis
 - We know which datasets were published as part of the thesis, and what publications they link to
 - UQ retains ownership of data after the student leaves

Q1 will see the release of further HDR process workflows into UQRDM

So...

Not all integration is the same

- More and more, 'DMP questions' are relevant questions for human/animal ethics or grant applications
- Whilst integrating with existing systems is a step in the right direction, it is not the end goal
- UQ is starting to embed research project lifecycle workflows directly into UQRDM
- Starting in 2025 with brand new animal and human ethics workflows

Practical benefits

For researchers

- **Huge decrease in manual data entry across multiple systems**
- **Integrated workflows, for example:**
 - Team member joins the research team. Lead investigator adds collaborator to their UQRDM record and is guided through the process:
 - “Will this user require access to storage?”
 - “Will this user require access to you Digital Research Notebook?”
 - “Will this user work with animals in your application 2024-AE00001?” -> create amendment
 - Team member off boarding (and notifications when team members leave)
 - Close out research projects, and all related activities in one workflow
- **Single source of outstanding actions, notifications**
- **Familiar (and trusted) interfaces**

Good evening Mr Jan Wisgerhof

7 active records | 0.00GB storage | 6 collaborators

From Dr Andrew Kneynsberg | Submitted Friday 20 September 2024 at 5:05 pm

You have a **new record** to approve

Record: draftcheck2

REVIEW RECORD

Practical benefits

For the university & professional staff

- **Increase in metadata completeness, maturity and accuracy**
- **Researchers will engage with the system more often as more workflows are embedded into one system. This provides more “up to date” metadata**
- **Cost savings for the University**
- **Increased oversight for professional staff. For example:**
 - When an ethics application approval is requested, ethics staff will be able to check whether the associated record has appropriate data storage
 - When data storage is requested for the record, provided options can be tailored (or limited) depending on ethics application answers (identifiable data, etc)

Thank you

Jan Wisgerhof

Senior Manager, Research Systems

j.wisgerhof@uq.edu.au

CRICOS 00025B • TEQSA PRV12080

