

# ARCHITECTING RESEARCH DATA EXPERIENCES FOR SCALE

Aligning Infrastructure Intelligence with Data  
Taxonomy in a Distributed, Resource-  
Constrained Research Landscape

**eResearch Australasia 2025**

**Ben Wu**

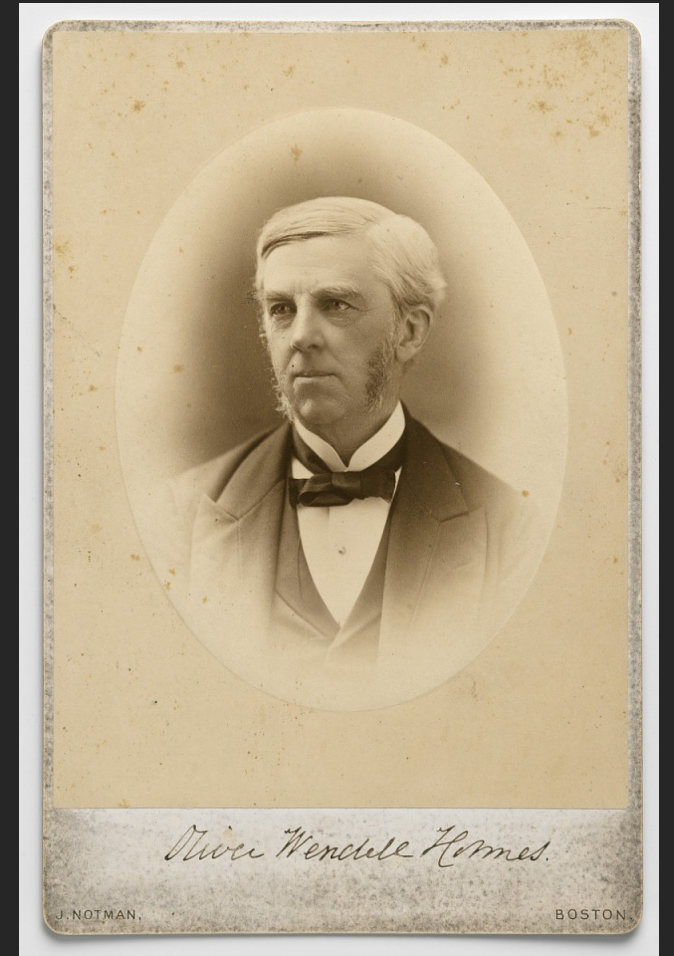
Principal Consultant, Business Consulting Services

21 OCTOBER 2025



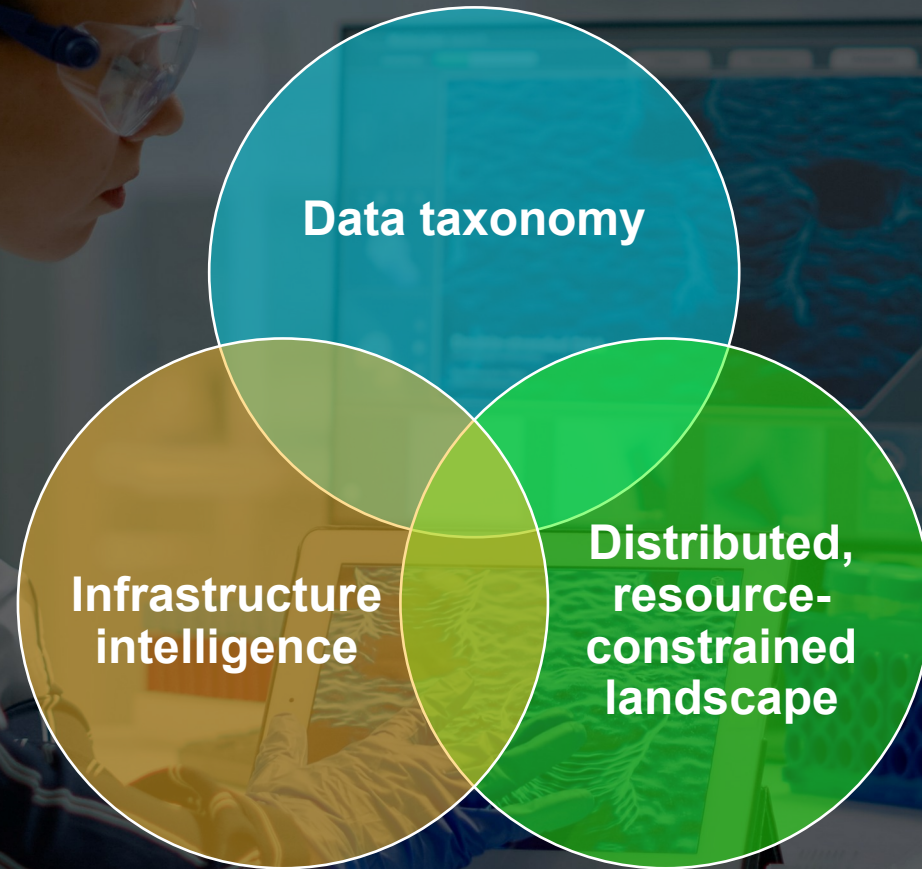
**“The mind, once expanded  
by a new idea, never  
returns to its original  
dimensions”**

Oliver Wendell Holmes Sr



# Architecting Research Data Experiences for Scale

Aligning Infrastructure Intelligence with Data Taxonomy in a Distributed, Resource-Constrained Research Landscape



# RESOURCE CONSTRAINTS - Decisions driven by resource constraints

**IMPERATIVE: deliver more for less!**

eResearch teams under pressure

Funding relief is temporary

**CSIRO researchers brace for hundreds of job cuts**

By **Joseph Brookes**  
Senior Reporter  
7 August 2025

The nation's research powerhouse has warned hundreds of jobs could be lost as the government slashes funding. The looming budget cuts are forcing CSIRO to restructure, with roles last year being reclassified as enterprise-level positions.

POWERED BY **HR LEADER** MOMENTUM MEDIA

“Look at the UOW’s books and you’ll see massive cuts. There was a \$39 million deficit, but their spending on international and domestic travel were enormous, and these are things you have to cut before you cut staff.”

“Anyone who’s run a business knows if you cut close to the bone, you have to reduce services, but University of Wollongong is claiming services will stay the same even after job cuts.”

“UOW management claiming ‘look we’ve saved 60 jobs’ but the number of people they’ll sack from 185 to 124, that’s what really upsets the community.”

This is just the latest in what has been a long string of budget cuts at universities in 2025, with Monash University, Deakin University, and the Australian National University all facing related issues.

## Australian National University announces more job cuts as it seeks to save \$250 million

By Adam Shirley  
Universities  
Thu 3 Jul



Amelia Dale is an English lecturer in the ANU School of Literature, Languages and Linguistics. (ABC News: Adam Shirley)

[abc.net.au/news/anu-announces-more-job-cuts-...](https://abc.net.au/news/anu-announces-more-job-cuts-...) Share

The Australian National University (ANU) is planning to cut another 59 jobs in its bid to save \$250 million by 2026. The additional job losses come on top of the 41 proposed job cuts announced last month.

Student caps lifted but still exist

Risk of wasting the opportunity to transform



**Faculty heads**

**Researchers**

**eRes data managers**

**Librarians**

**Data infrastructure teams**

**Everyone!**

# DISTRIBUTED - The Future of Research Isn't Tied to a Campus

Research is no longer bound by walls or borders

Australian institutions need to grow “pop-up campuses” and virtual collaborations around the world

Every jurisdiction brings unique data laws, ethics rules, and infrastructure challenges

# DATA TAXONOMY - The Cost of Incoherence

Each country has its own privacy laws, retention rules, and storage constraints

Every project team seems to reinvent the wheel: different systems, taxonomies, and governance

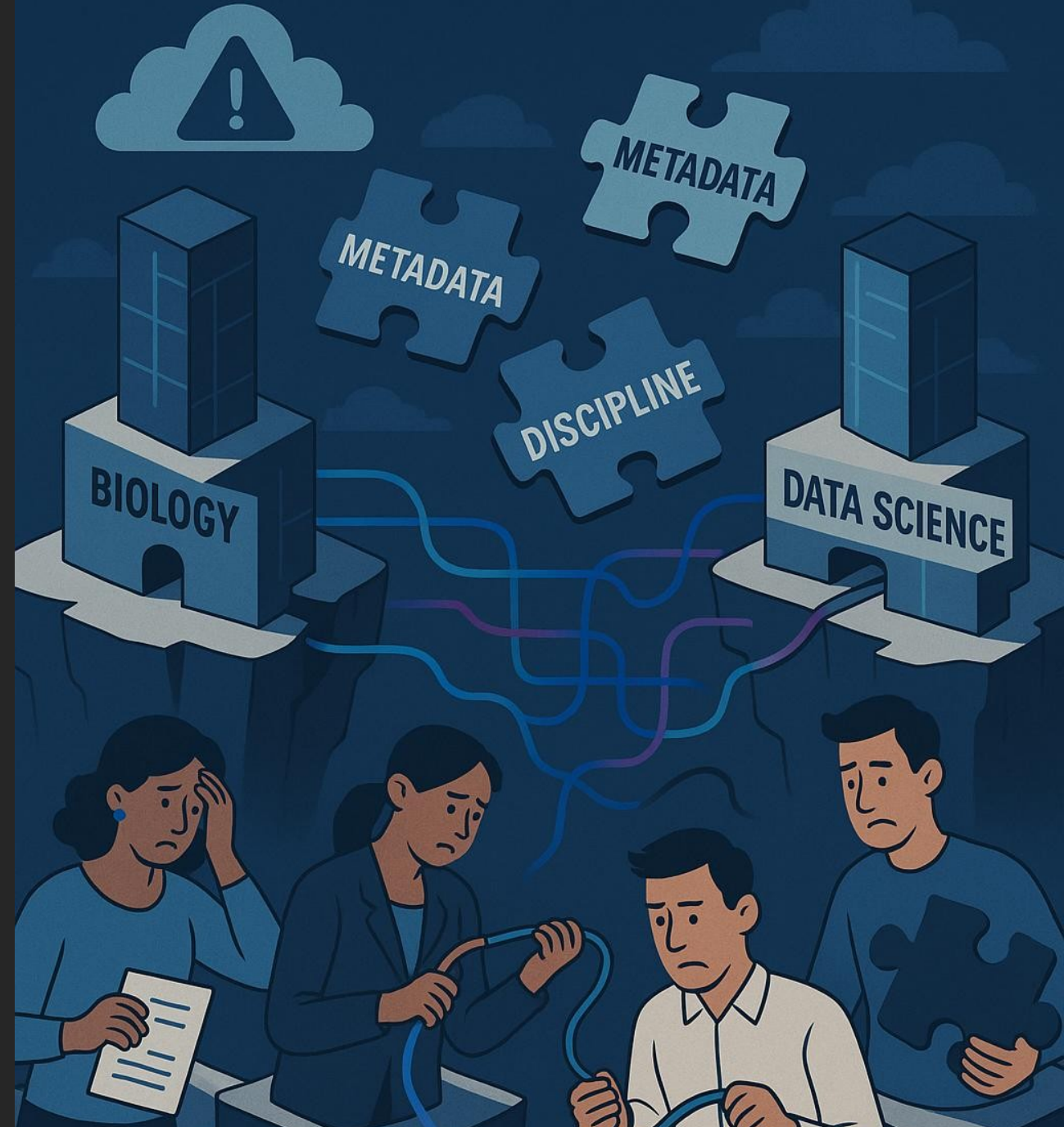
The result?  
Friction, cost, and loss of trust

# DATA TAXONOMY & the RDM challenge

Fragmentation, Lineage, and Provenance Gaps

*“Despite guidelines that require researchers and institutions to manage the retention or disposal of research generated content, such management decisions rarely occur resulting in an accumulation of poorly understood and rapidly growing data holdings, some of which might be valuable, and some of which might be digital research debris”*

Business Intelligence and Reporting of Research Data, Nov 2023, DOI: 10.5281/zenodo.10076883

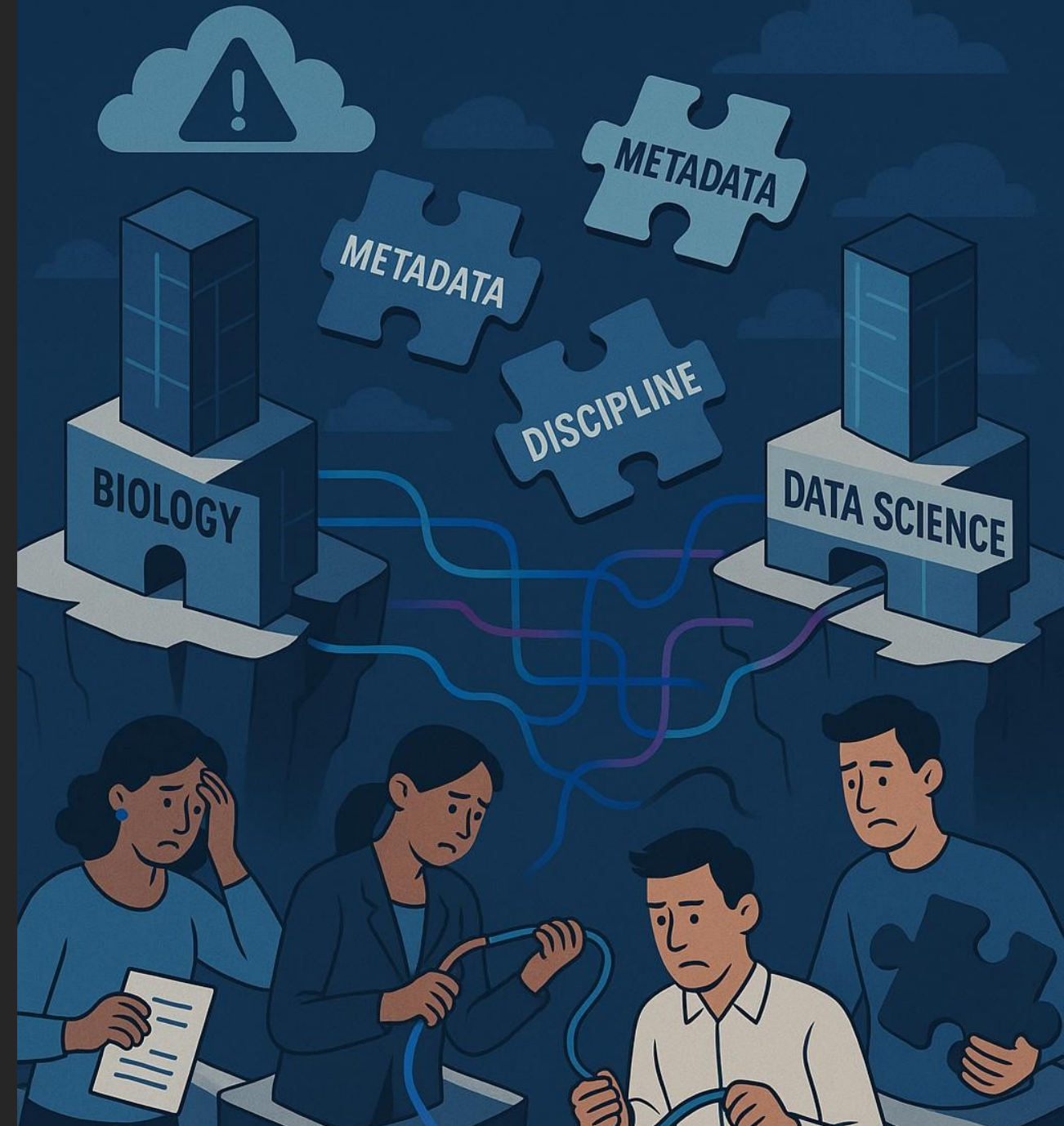


# DATA TAXONOMY & the RDM challenge

## Fragmentation, Lineage, and Provenance Gaps

- Siloed systems
- Inconsistent lineage & provenance
- Weak integration
- Taxonomy and metadata inconsistencies
- Researcher adoption hindered

**These factors inhibit effective research data management at scale**



# Frameworks, Taxonomies & Collaboration

## National and Global Alignment in Practice

- **ARDC National eRDM Framework**

10 primary and 9 secondary elements

- **RDA global standards**

Provenance, metadata, data citation

- **QCIF's ReDBox**

Aligning local RDM to national metadata schemas

- **Taxonomies = “digital Rosetta Stone”**

Interoperability, inter-usability

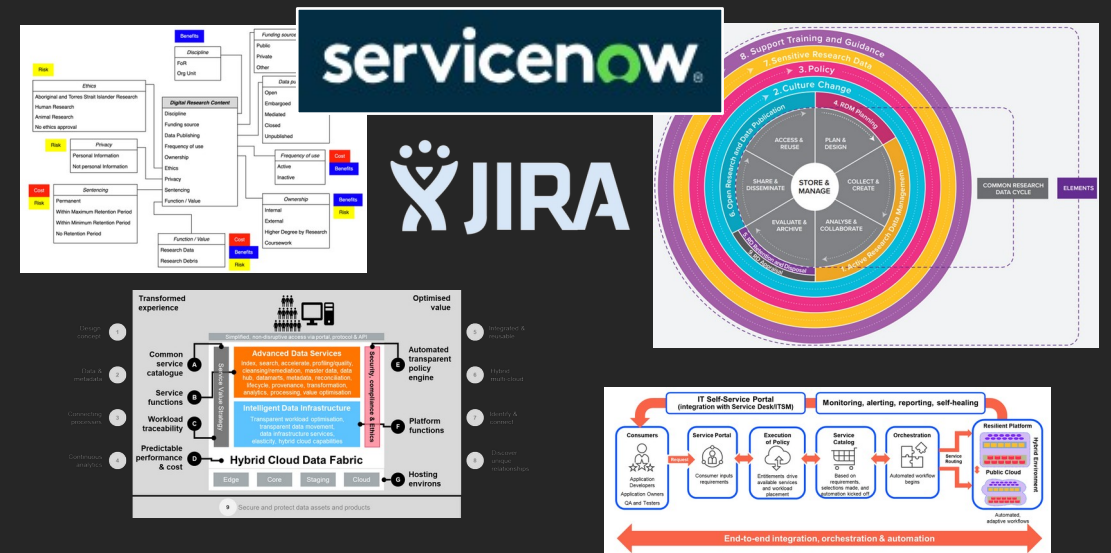
- **Open-source tools**

Shared backbone Kafka, Postgres, Clickhouse, OpenSearch, OpenMetadata

- **Intelligent Data Infrastructure**

Composable, scalable, policy-driven, cost optimised, cyber-resilient

# Metadata pools with immature integration & orchestration

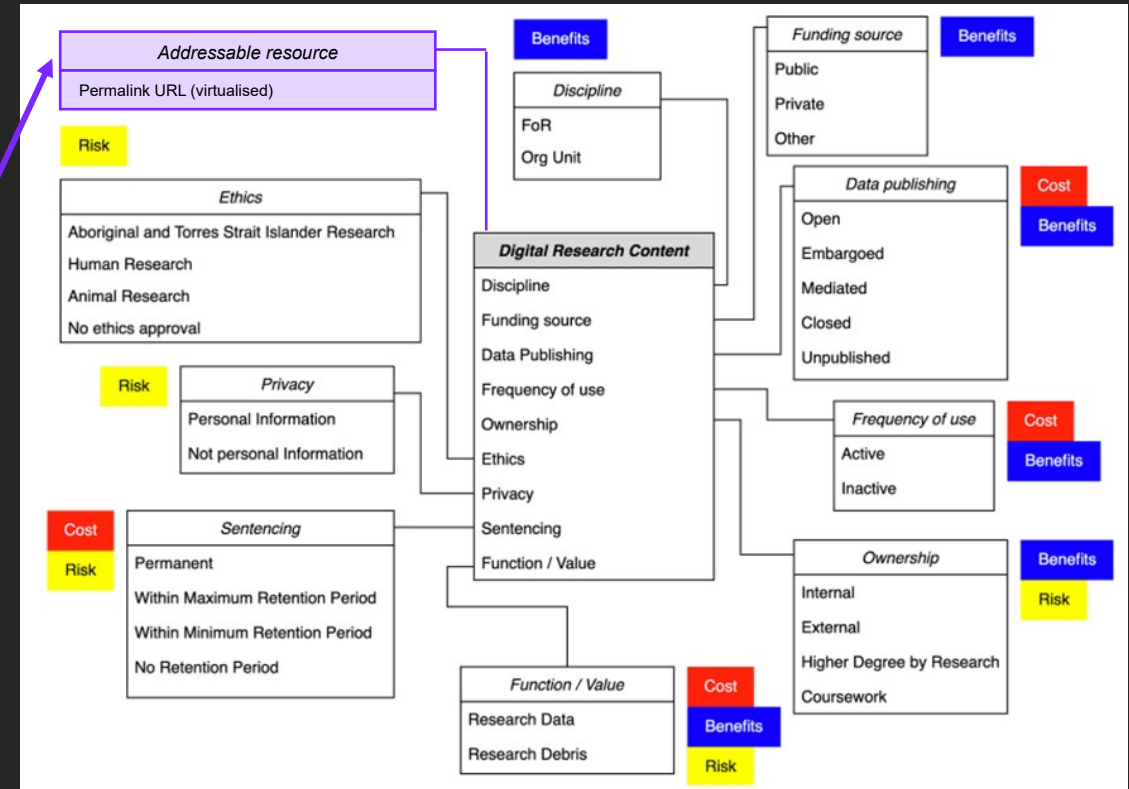


# Research Data Taxonomy – The Common Language

Fix the way we describe and map research data before we scale globally

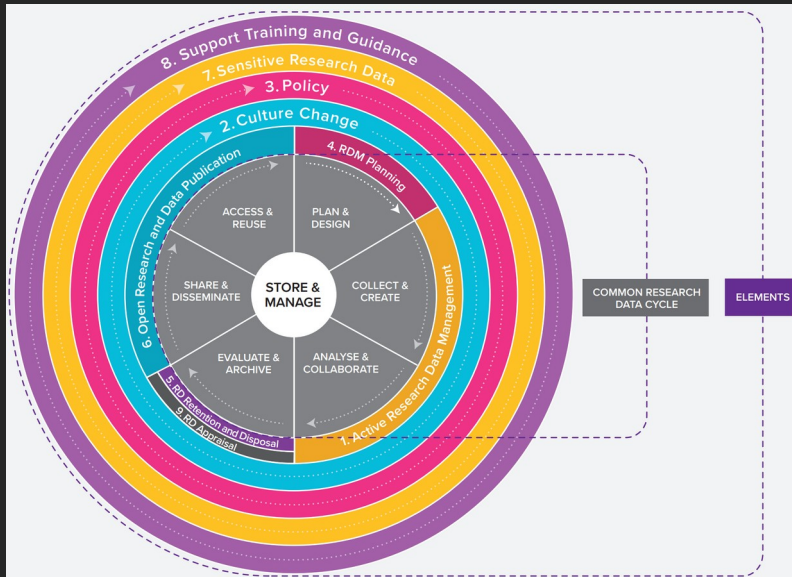
- A national (and globally mappable) taxonomy gives research data shared meaning
- It defines not just what the data is, but how it must be treated—across domains and borders
- Enables automated mapping between national and international schemas

An addressable resource reference (permalink URL) accelerates research by making data machine identifiable and readable

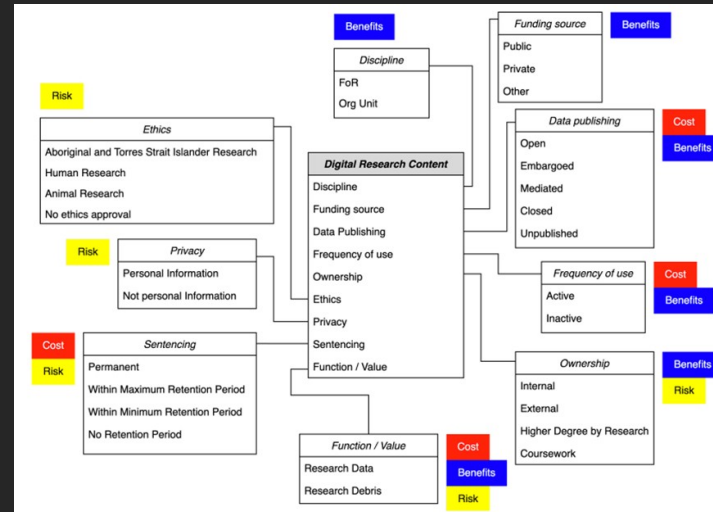


# Research data taxonomy – bringing together user needs and infrastructure

## eResearch Data Lifecycle



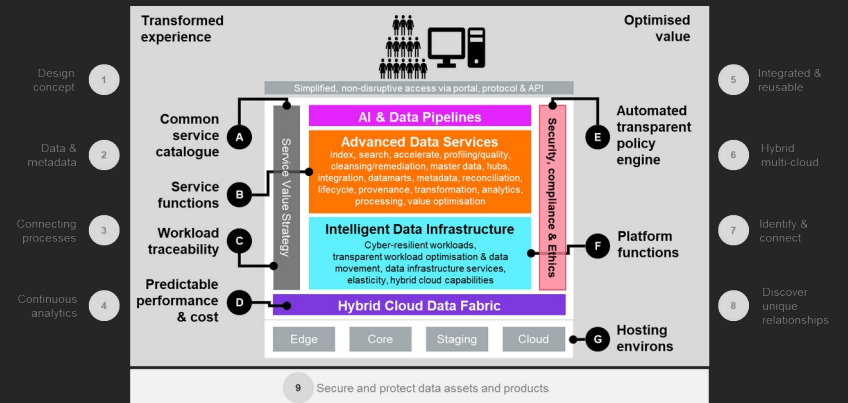
## Taxonomy needed for linkage



What

How

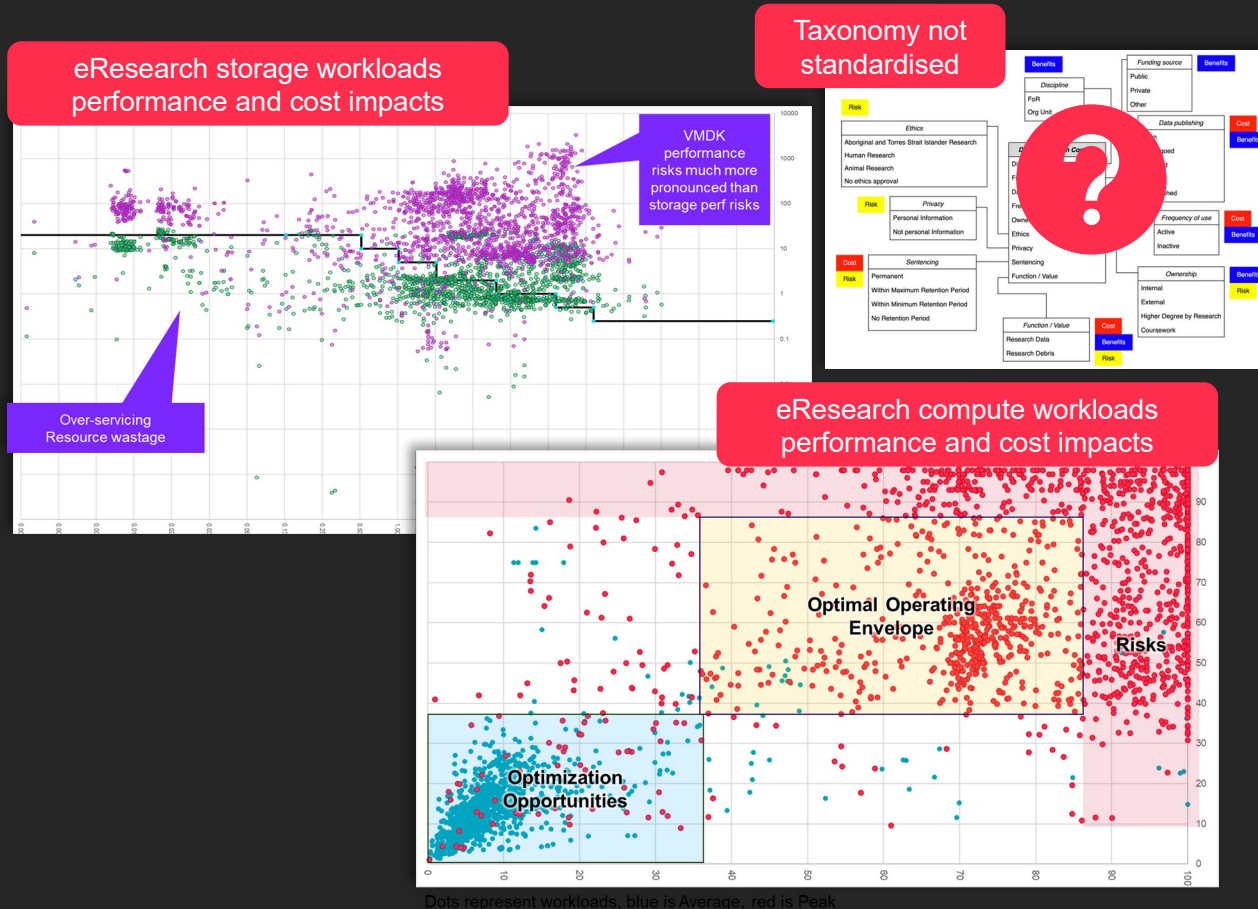
## eResearch Data Infrastructure



Australian Research Data Commons. (2023) Research Data Management Framework for Institutions. Viewed online at: <http://doi.org/10.5281/zenodo.6392340>

# eResearch IT is still a hot mess!

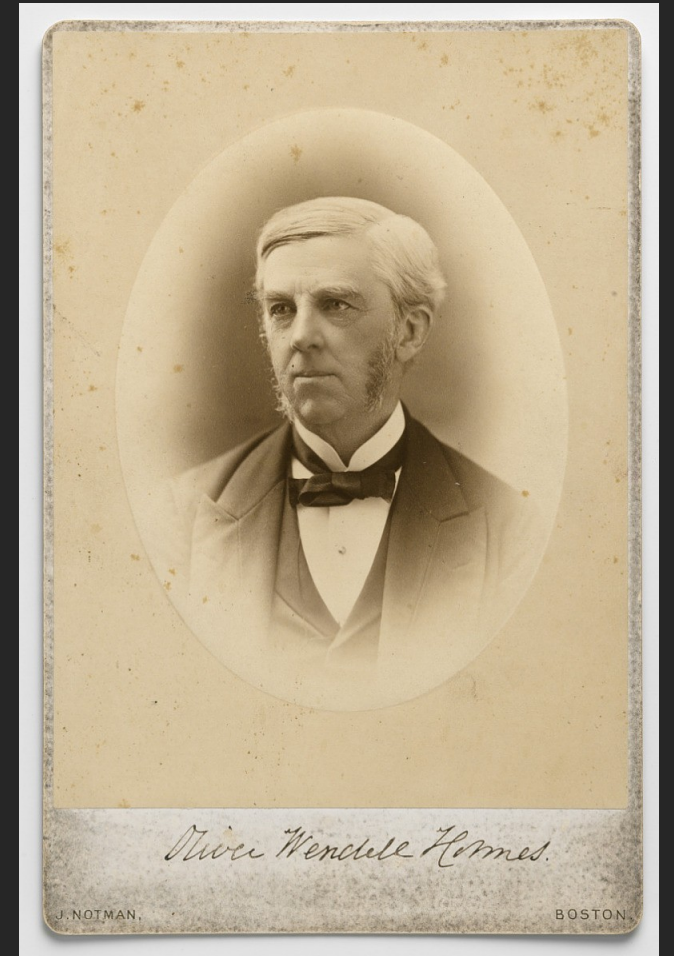
Case study – Academic organisation with global campuses, yet severely impacted by funding caps



- Compute and storage workloads are not balanced between cost and performance
- Metadata not in place to link research assets to underlying storage
- Adhoc descriptors, not all data aligned with research taxonomy
- No scalable architecture to address data for diverse compliance requirements
- No safe, efficient way to share research data

**“The mind, once expanded  
by a new idea, never  
returns to its original  
dimensions”**

Oliver Wendell Holmes Sr



# **ACTION #1**

**GET OUR**

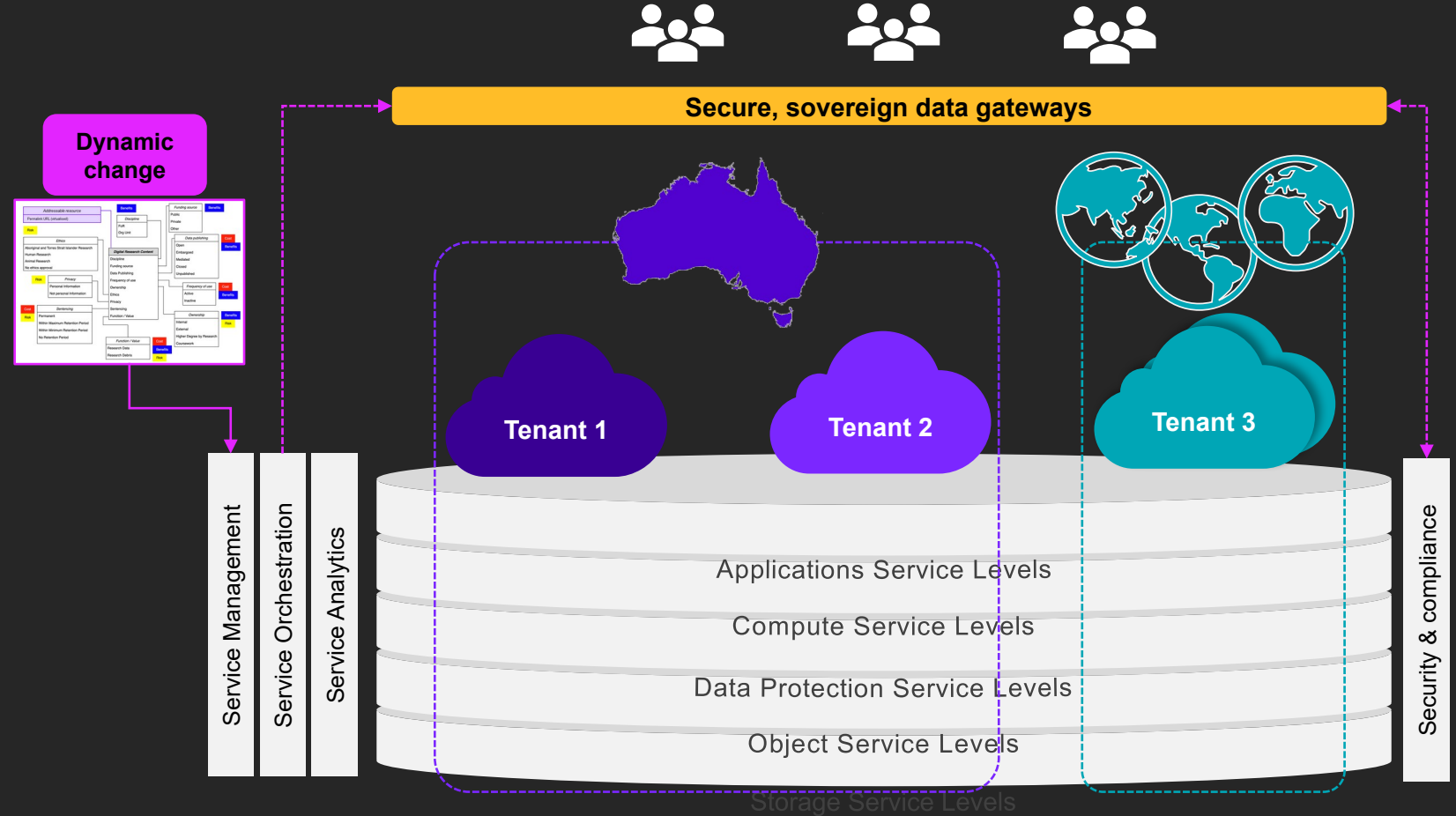
**INTELLIGENT INFRASTRUCTURE  
REFERENCE MODELS**

**AT NO COST**

# Architecting Research Data Experiences for Scale

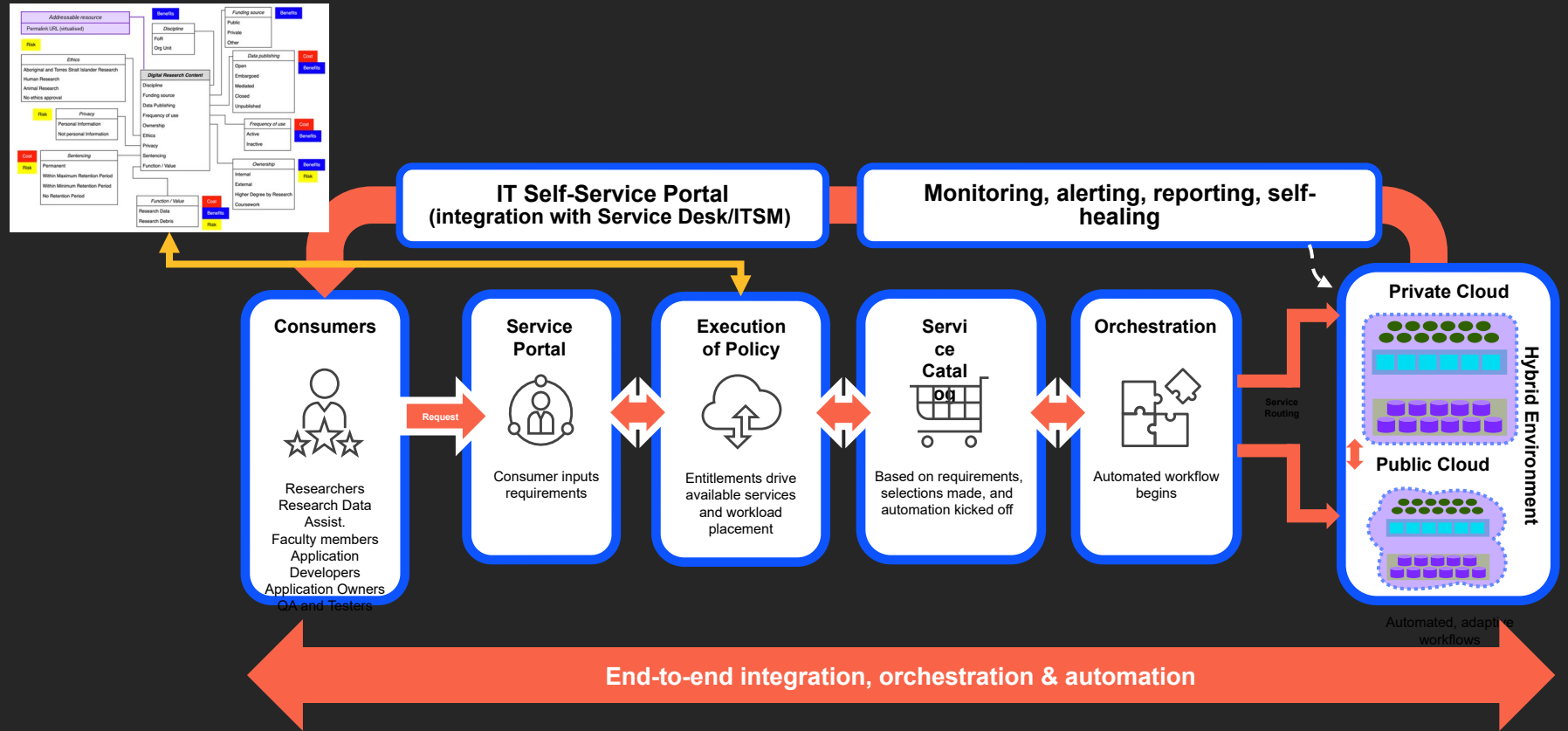
Intelligent infrastructure responds dynamically to changes in data attributes using approved policies

- Researchers get consistent access and collaboration experiences anywhere
- Infrastructure teams deploy new environments repeatably and securely
- Virtual and physical implementations functionally indistinguishable
- Institutions maintain sovereignty and compliance while enabling rapid innovation



# Intelligent Infrastructure – Making Compliance Automatic

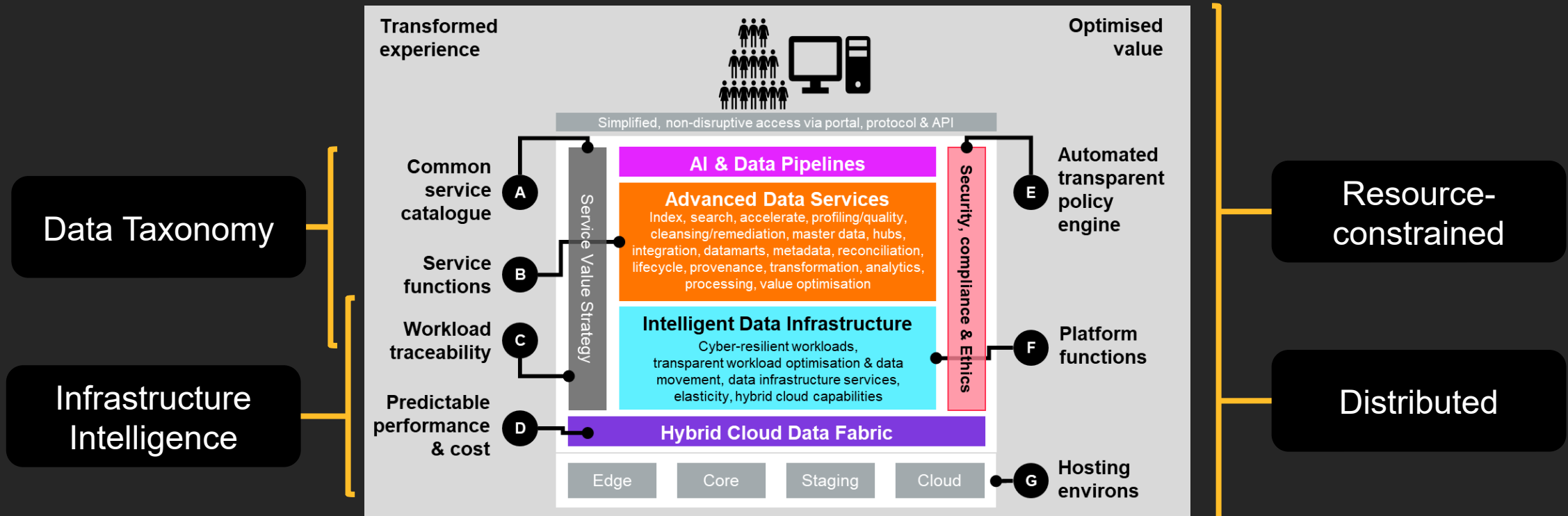
- Infrastructure aligned to taxonomy enforces governance automatically
- Policies follow data as it moves between institutions and countries
- Multi-tenant, hybrid-cloud architectures enable rapid deployment of new research nodes



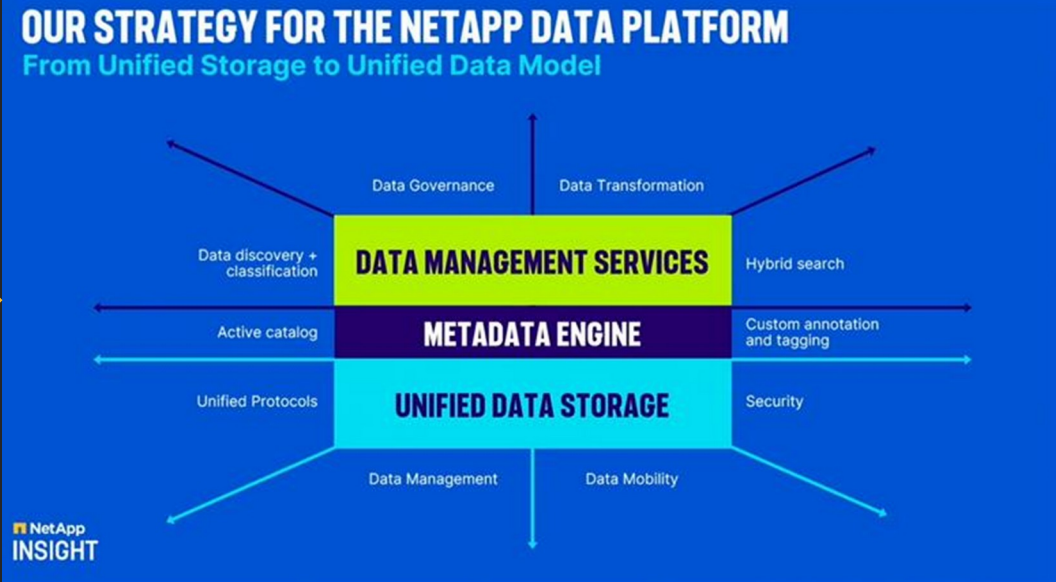
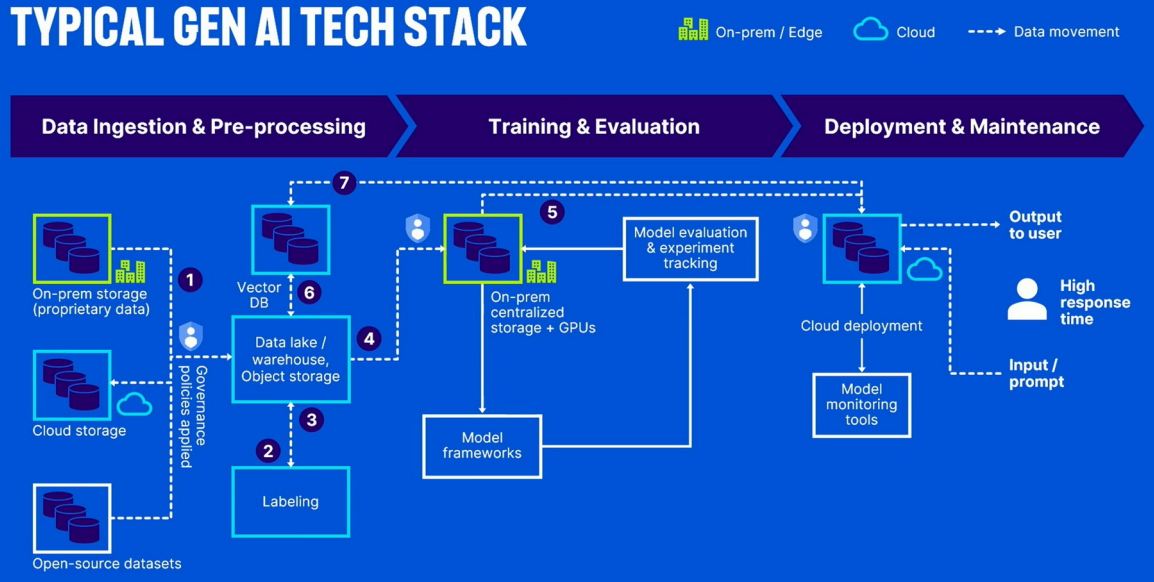
# Architecting Research Data Experiences for Scale

## Aligning Infrastructure Intelligence with Data Taxonomy in a Distributed, Resource-Constrained Research Landscape

### A New Approach for Scalable, Responsible Research



# New research infrastructure architectures lead to improved outcomes



# ACTION #2

REQUEST AN

ERESEARCH WORKLOAD ROI  
OPTIMISATION ANALYSIS

AT NO COST\*

\* subject to qualification

# Review, identify and optimise workload performance, minimising risk and cost

Improving eResearch experiences and recovering significant budget for reinvestment in the future



Resource constraints are our reality. But we must still optimise the value of eResearch investments

# **ACTION #3**

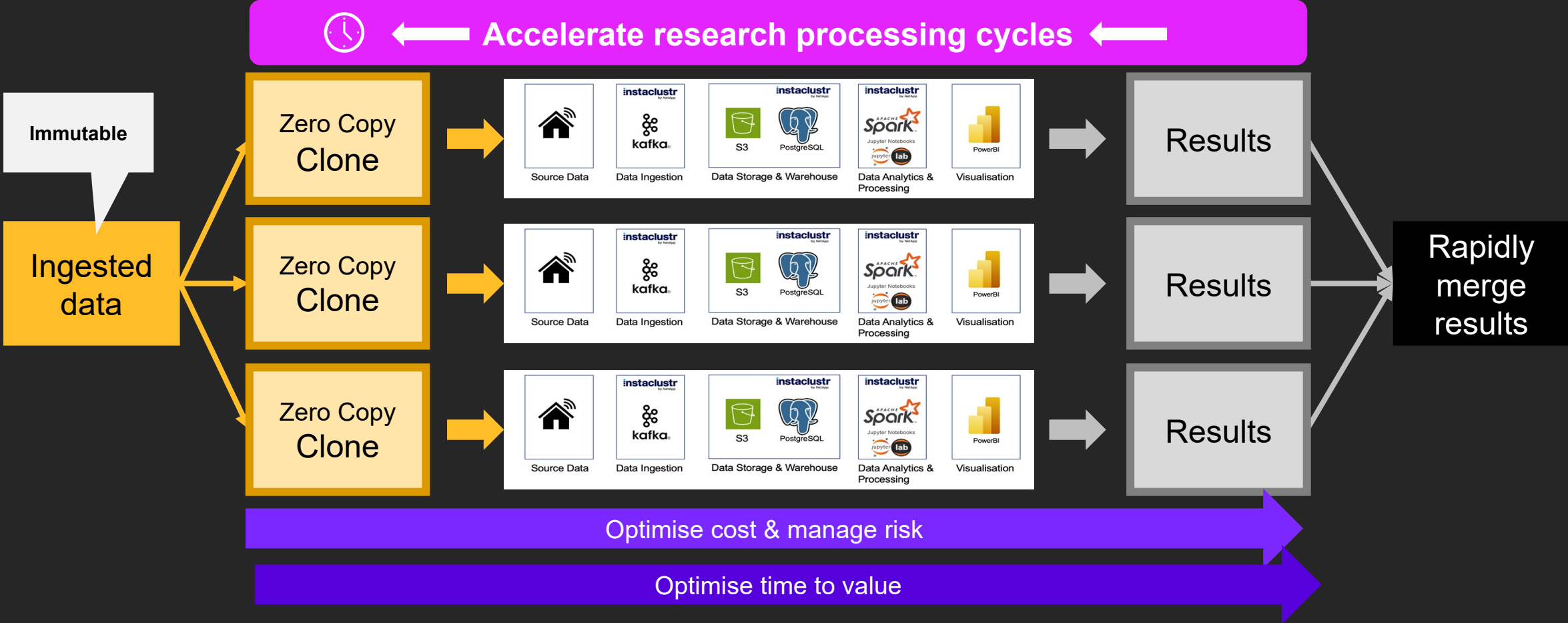
**TEST DRIVE**

**INTELLIGENT WORKLOAD CLONING  
FOR AI & ERESEARCH**

**AT NO COST**

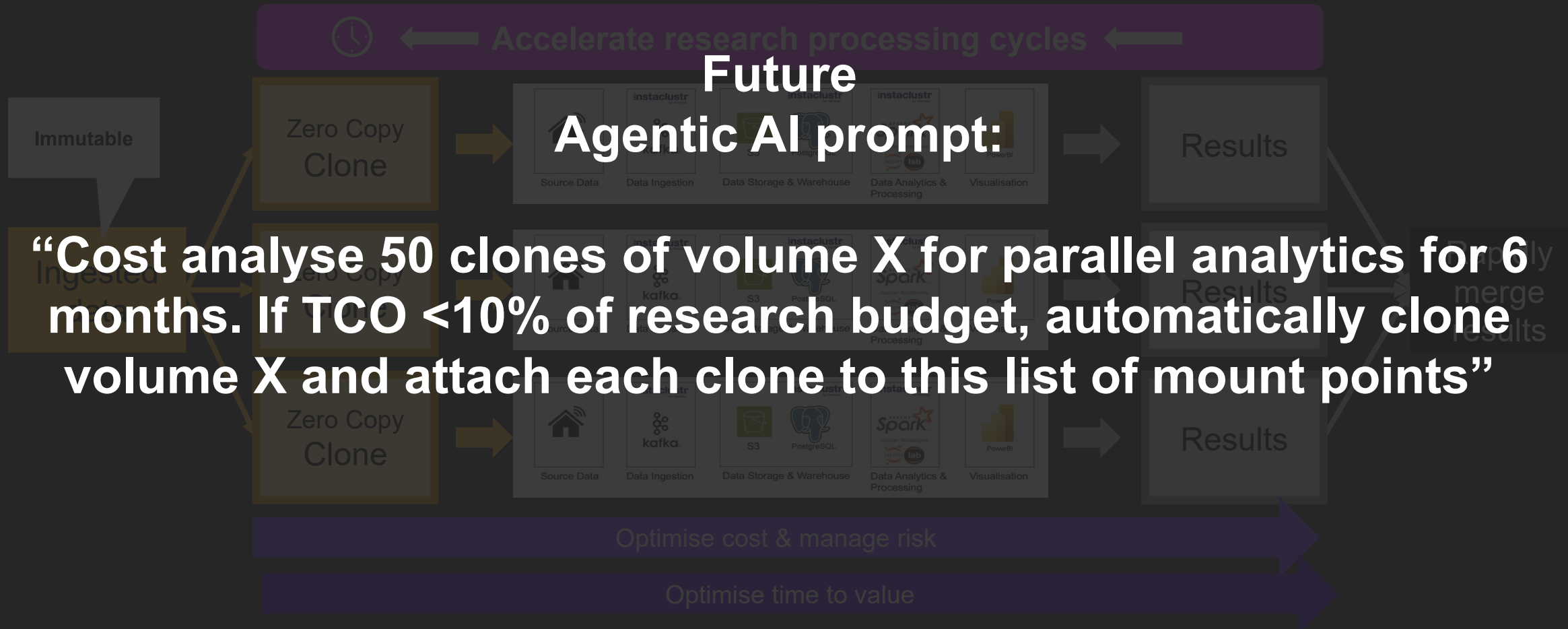
# Low-hanging fruit for eResearch Data – data cloning you can explore today!

Generate, track, trace and delete data clones safely by policy – effectively govern  
Rethink data strategy – extreme agility, efficiency, cost-effectiveness



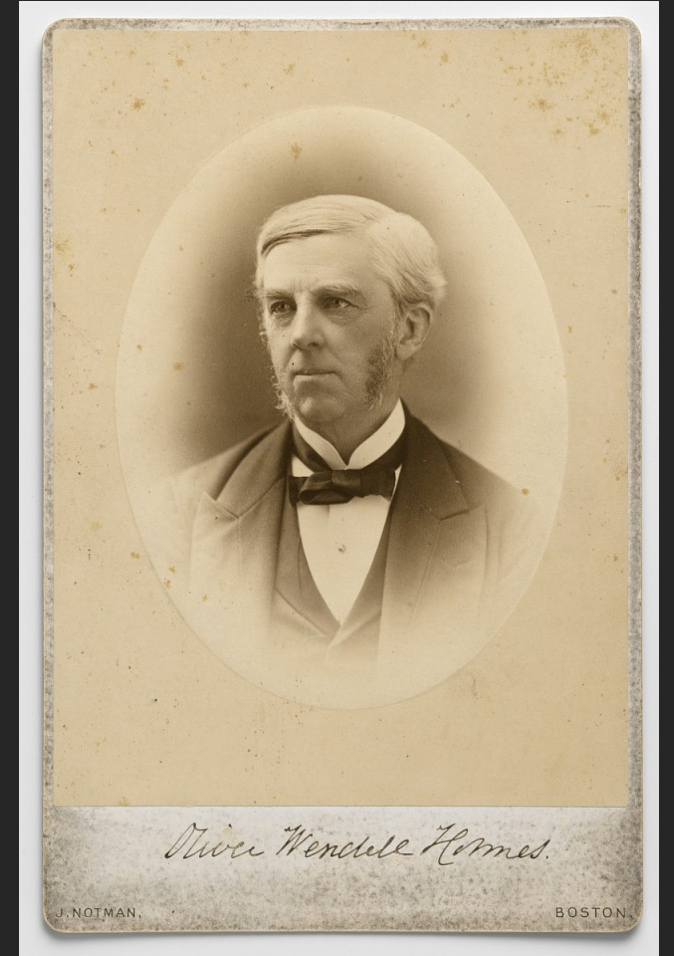
# Low-hanging fruit for eResearch Data – data cloning you can do today!

Generate, track, trace and delete data clones safely by policy – effectively govern  
Rethink data strategy – extreme agility, efficiency, cost-effectiveness



**“The mind, once expanded  
by a new idea, never  
returns to its original  
dimensions”**

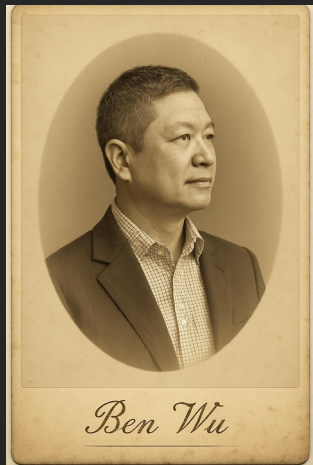
Oliver Wendell Holmes Sr



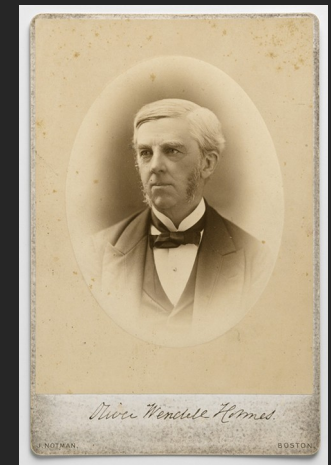
# **DATA TAXONOMY INTELLIGENT INFRASTRUCTURE DISTRIBUTED, RESOURCE-CONSTRAINED**

**ACTION 1: GET THE REFERENCE MODEL**  
**ACTION 2: ANALYSE YOUR WORKLOADS**  
**ACTION 3: EXPLORE DATASET CLONING**

# NetApp eResearch Data Workshops



**Address data challenges  
Innovate with new technologies  
Deliver improved value**



**[benjamin.wu@netapp.com](mailto:benjamin.wu@netapp.com)**



**THANK YOU**

**[benjamin.wu@netapp.com](mailto:benjamin.wu@netapp.com)**