

Establishing the foundations for a research organisation to leverage AI in animal science

Mark Neal, Jenny Jago

AI4AS

4 June 2025

Dairynz 

Our Purpose

Progressing a positive future for New Zealand dairy farming

Our Vision

To make the levy the best investment of every New Zealand dairy farmer

Our strategic priorities



Accelerating on-farm productivity

World-leading animal genetic gain

Improved forage gains

Increased workplace productivity



Powering more adaptable & resilient farms

Future-fit farm systems

Credible evidence

Strong biosecurity systems



Enabling sustainable & competitive dairying

Healthy waterways

Enhanced animal care

Reduced GHG emissions

Our key activities

With farmers, we deliver:

A comprehensive foundation of farming knowledge & solutions

Insightful & targeted benchmarking

High-impact partnerships

Meaningful connections

Progress research & science

Provide industry analytics & insight

Inform pragmatic, evidence-based policy

Develop targeted solutions

Lead collective action

For the dairy sector we:

The DairyNZ difference

We are a preferred partner because we are independent and represent all dairy farmers

We have world-class expertise relevant to our complex industry challenges

We take a farm systems approach in how we work with farmers

Powered by our people: One team at DairyNZ, working for every New Zealand dairy farmer

Our Purpose

Progressing a positive future for New Zealand dairy farming

Our Vision

To make the levy the best investment of every New Zealand dairy farmer

Our strategic priorities



Accelerating on-farm productivity

World-leading animal genetic gain

Improved forage gains

Increased workplace productivity



Powering more adaptable & resilient farms

Future-fit farm systems

Credible evidence

Strong biosecurity systems



Enabling sustainable & competitive dairying

Healthy waterways

Enhanced animal care

Reduced GHG emissions

Our key activities

With farmers, we deliver:

A comprehensive foundation of farming knowledge & solutions

Insightful & targeted benchmarking

High-impact partnerships

Meaningful connections

For the dairy sector we:

Progress research & science

Provide industry analytics & insight

Inform pragmatic, evidence-based policy

Develop targeted solutions

Lead collective action

The DairyNZ difference

We are a preferred partner because we are independent and represent all dairy farmers

We have world-class expertise relevant to our complex industry challenges

We take a farm systems approach in how we work with farmers

Powered by our people: One team at DairyNZ, working for every New Zealand dairy farmer

Our Purpose

Progressing a positive future for New Zealand dairy farming

Our Vision

To make the levy the best investment of every New Zealand dairy farmer

Our strategic priorities



**Accelerating
on-farm productivity**



**Powering more adaptable
& resilient farms**



**Enabling sustainable
& competitive dairying**

World-leading
animal
genetic gain

Improved
forage gains

Increased
workplace
productivity

Future-fit farm
systems

Credible
evidence

Strong
biosecurity
systems

Healthy
waterways

Enhanced
animal care

Reduced GHG
emissions

Our key activities

With farmers, we deliver:

A comprehensive
foundation
of farming
knowledge
& solutions

Insightful
& targeted
benchmarking

High-impact
partnerships

Meaningful
connections

Progress
research &
science

Provide
industry
analytics
& insight

Inform
pragmatic,
evidence-
based policy

Develop
targeted
solutions

Lead
collective
action

The DairyNZ difference

**We are a preferred partner
because we are independent
and represent all dairy farmers**

**We have world-class expertise
relevant to our complex
industry challenges**

**We take a farm systems
approach in how we
work with farmers**

Powered by our people: One team at DairyNZ, working for every New Zealand dairy farmer

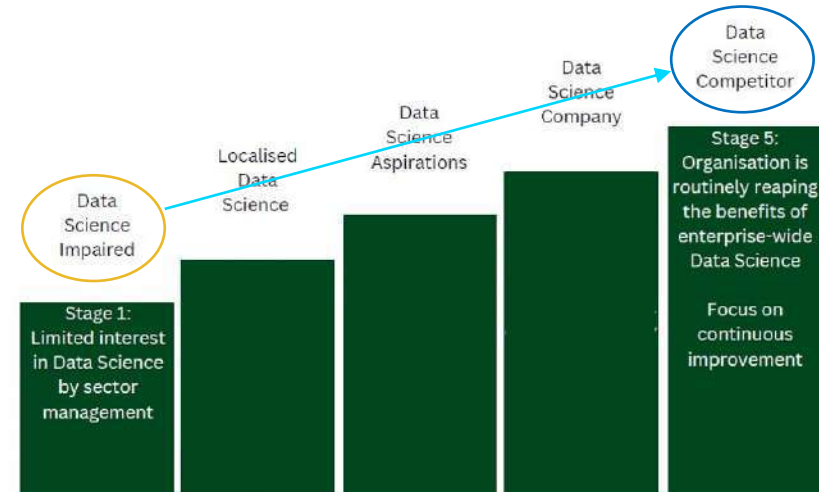
Drivers – Modern science is data science

- Range and volume of data
- Modern workflows (code and data) and techniques (Machine Learning, AI)
- Minimise risks
- Quality, credibility and impact



Drivers – Modern science is data science

- Range and volume of data
- Modern workflows (code and data) and techniques (Machine Learning, AI)
- Minimise risks
- Quality, credibility and impact



Step 1: Pitch the Plan



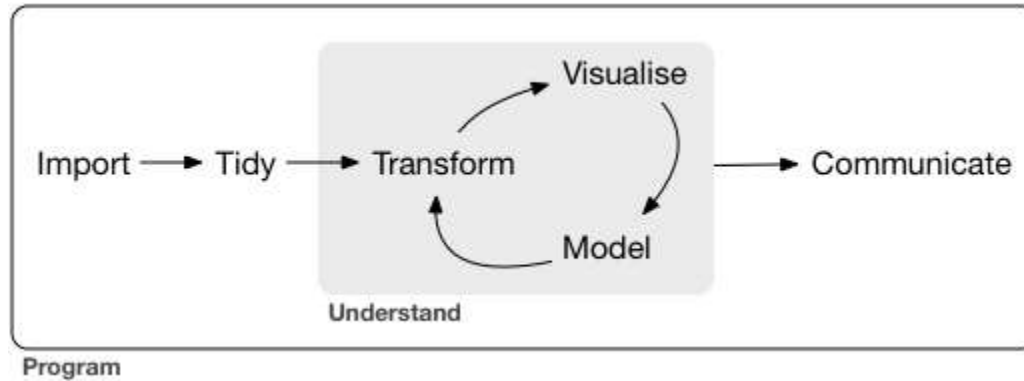
What needed to happen?

- Multi-year vision
- Modern Science Workflows project
 - Infrastructure
 - Skills and Capability
 - Business disciplines

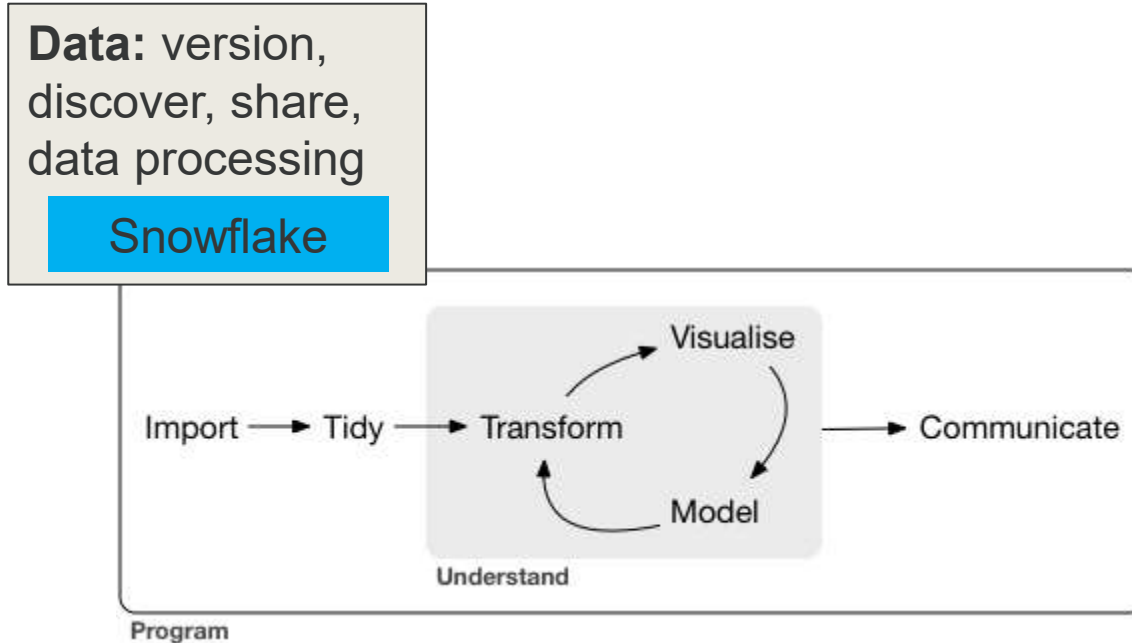
Step 2: Plan Into Action



Infrastructure



Modern Science Workflow Infrastructure



Snowflake (Data warehouse and compute)

- Data ingestion and connections
 - Bronze -> Silver -> Gold
- Efficient compute
 - Relative to local laptop
- Governance plan
 - Accounts for Access

```
> toc()
9 035HOLHOL840M003135762776 54.7
10 035HOLHOL840M003135874925 58.9
# ... with more rows
> toc()
dplyr snowflake: 0.89 sec elapsed
>
> tic("dplyr local")
> results <- raw_holconf_long %>%
+   group_by(bull_int_id) %>%
+   summarise(Avg_BV = mean(V12, na.rm = TRUE))
# summarise() ungrouping output (override with
> results
# A tibble: 192,862 x 2
   bull_int_id      Avg_BV
   <chr>          <dbl>
1 035HOLBRFGBRM100756401896 42.5
2 035HOLBRFGBRM100756501897 42.5
3 035HOLBRFGBRM161594701047 42.2
4 035HOLBRFGBRM171632301729 41.0
5 035HOLBRFGBRM200910102057 43.5
6 035HOLBRFGBRM200910102148 41.6
7 035HOLBRFGBRM200910202058 43.5
8 035HOLBRFGBRM200910701846 43.1
9 035HOLBRFGBRM200924101784 42.9
10 035HOLBRFGBRM201323603588 43.2
# ... with 192,852 more rows
> #as.data.frame(results)
> toc()
dplyr local: 3.47 sec elapsed
```

Central ingestion, catalogue data

Backlog	Discovered	Azure	Snowflake	Refresh
<div><div><div><div><div><div></div><div>New item</div></div><div></div></div><div><div>5594 Research Technician data (e.g. ovens)</div><div>State New</div></div></div><div><div>5593 Geospatial data pipeline, recording and upload of technician data (e.g. calibration cuts)</div><div>State New</div></div></div><div><div>5591 Static Liveweight Data</div><div>State New</div></div><div><div>5589 Walk over Weigh prototype data</div><div><div>Mark Neal</div><div>State New</div></div></div><div><div>5306 Sustainable Food & Fibres (SFF)</div><div><div>Roshean Woods</div><div>State New</div></div></div><div><div>5590 BCS Camera prototype Data</div><div>State New</div></div><div><div>5321 Telford Datasets</div><div>State New</div></div></div>	<div><div>5301 Puberty Scale Up 1</div><div><div>Robbie McGregor</div><div>State Active</div><div>Chris Burke Susan Meier</div></div></div> <div><div>5308 LowN</div><div><div>Robbie McGregor</div><div>State Active</div></div></div> <div><div>5322 DairyBase</div><div><div>Robbie McGregor</div><div>State Active</div><div>Mark Neal</div></div></div>	<div><div>5298 CSV Farm Data</div><div><div>Mickey Lulu</div><div>State Active</div><div>Charlotte Reed</div></div></div> <div><div>5297 Aii Milk</div><div><div>Mickey Lulu</div><div>State Active</div><div>Charlotte Reed</div></div></div> <div><div>5595 Sleep Studies</div><div><div>Robbie McGregor</div><div>State Active</div><div>Collum</div></div></div>	<div><div>4107 NS&C: Platimeters Sync</div><div><div>Robbie McGregor</div><div>State Active</div><div>NS&C</div></div></div> <div><div>5309 Greenhouse Gas Inventory Project</div><div><div>Mickey Lulu</div><div>State Active</div><div>Mark Neal</div></div></div> <div><div>5318 Virtual Climate Station</div><div><div>Robbie McGregor</div><div>State Active</div><div>Hemds Levy</div></div></div> <div><div>5588 Animal Care Consult Database</div><div><div>Robbie McGregor</div><div>State Active</div></div></div> <div><div>5324 GPSSAT</div><div><div>Robbie McGregor</div><div>State Active</div><div>Wayne Hoffman</div></div></div>	<div><div>5299 Smaxtec</div><div><div>Josh Barrett</div><div>State Active</div><div>Charlotte Reed NZBIDA Paul Edwards</div></div></div> <div><div>5323 QCONZ</div><div><div>Josh Barrett</div><div>State Active</div><div>Wayne Hoffman</div></div></div> <div><div>5296 WeatherLink</div><div><div>Josh Barrett</div><div>State Active</div></div></div> <div><div>4106 NS&C: AgTech360 Data</div><div><div>Josh Barrett</div><div>State Active</div><div>NS&C</div></div></div> <div><div>5307 Frontier Farms</div><div><div>Robbie McGregor</div><div>State Active</div></div></div>



FAIR principles

Discovery with MS Purview

The screenshot displays the Microsoft Purview web interface. The top navigation bar includes the Microsoft Purview logo, a search bar, and icons for Copilot, notifications, and settings. The left sidebar contains navigation links for Home, Unified Catalog, Overview, Discovery, Catalog management, Health management, and Related solutions (Data Map). The main content area shows search results for 'plantain'. A filters panel on the left lists categories like Asset Type, Data source type, Collection, Classification, and Contact. The results list includes three items: 'PARTNER_FARM_PLANTAIN_MONITORING' (Snowflake Table), 'STACKEDFARMLET PASTURE BOTANICAL COMPOSITION Copy' (Azure Blob Resource Set), and 'Taratua Plantain' (Power BI Dataset). Each result shows its name, icon, description, fully qualified name, and update date.

Microsoft Purview

Search

Copilot

Home

Solutions

Learn

Settings

Unified Catalog

Related solutions

Data Map

Unified Catalog

Overview

Discovery

Catalog management

Health management

Search results for plantain

Search: plantain

Filter

Filters

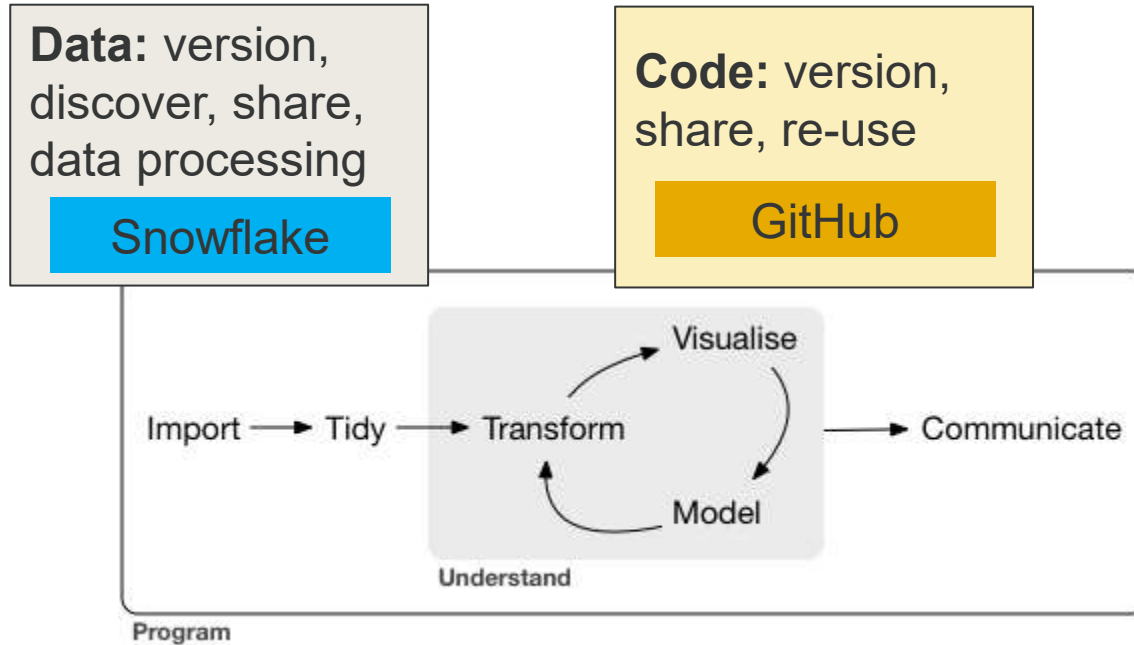
- Asset Type
 - Data
 - Business
 - Governance
 - Organization
- Data source type
- Collection
- Classification
- Contact

Showing 180+ results

Sort by: Relevance

- ☐ **PARTNER_FARM_PLANTAIN_MONITORING** Snowflake Table
No description available
Fully qualified name: snowflake://dnz_economics.australia-east.azure.snowflakecomputing.com/databases/...
Updated 3 months ago
- ☐ **STACKEDFARMLET PASTURE BOTANICAL COMPOSITION Copy** Azure Blob Resource Set
No description available
Fully qualified name: https://sadalakeue.blob.core.windows.net/econstaging/DataServices/lowN/stacke...
Updated 8 days ago
- ☐ **Taratua Plantain** Power BI Dataset
No description available
Fully qualified name: https://app.powerbi.com/groups/5dfb212e-eaa7-48f0-8d5e-857fc12a490/datasets/9...
Updated a year ago
- ☐ **Taratua Plantain** Power BI Report

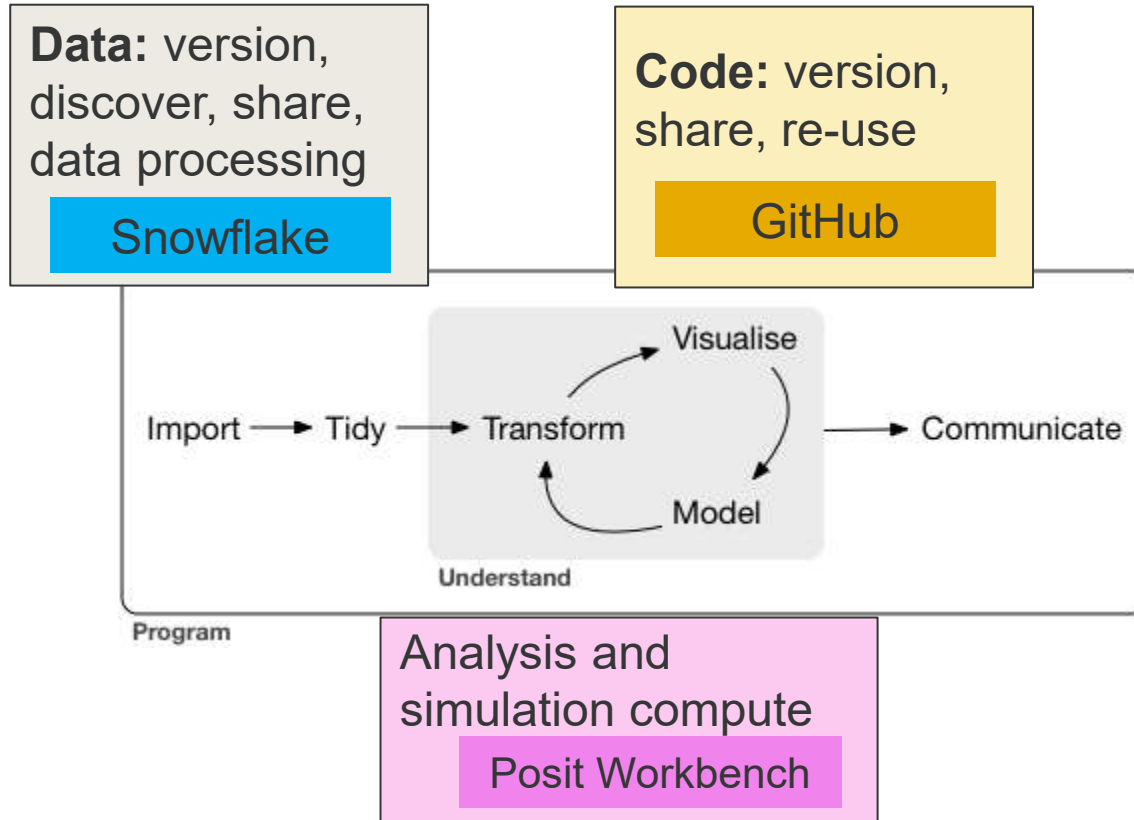
Modern Science Workflow Infrastructure



GitHub

- Account for organisation
- Expectations
 - What lives where?
 - Discoverability and reuse
 - Best practice password storage

Modern Science Workflow Infrastructure

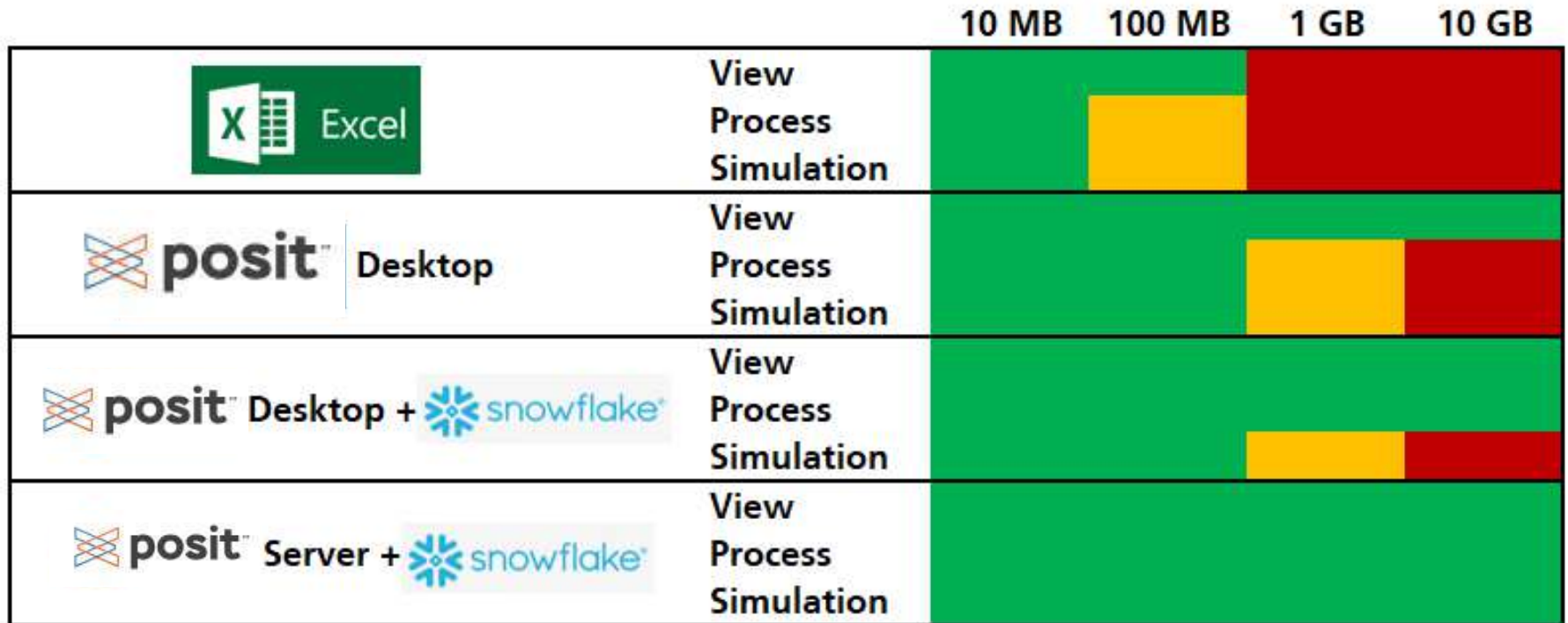


Posit Workbench (Analysis compute)

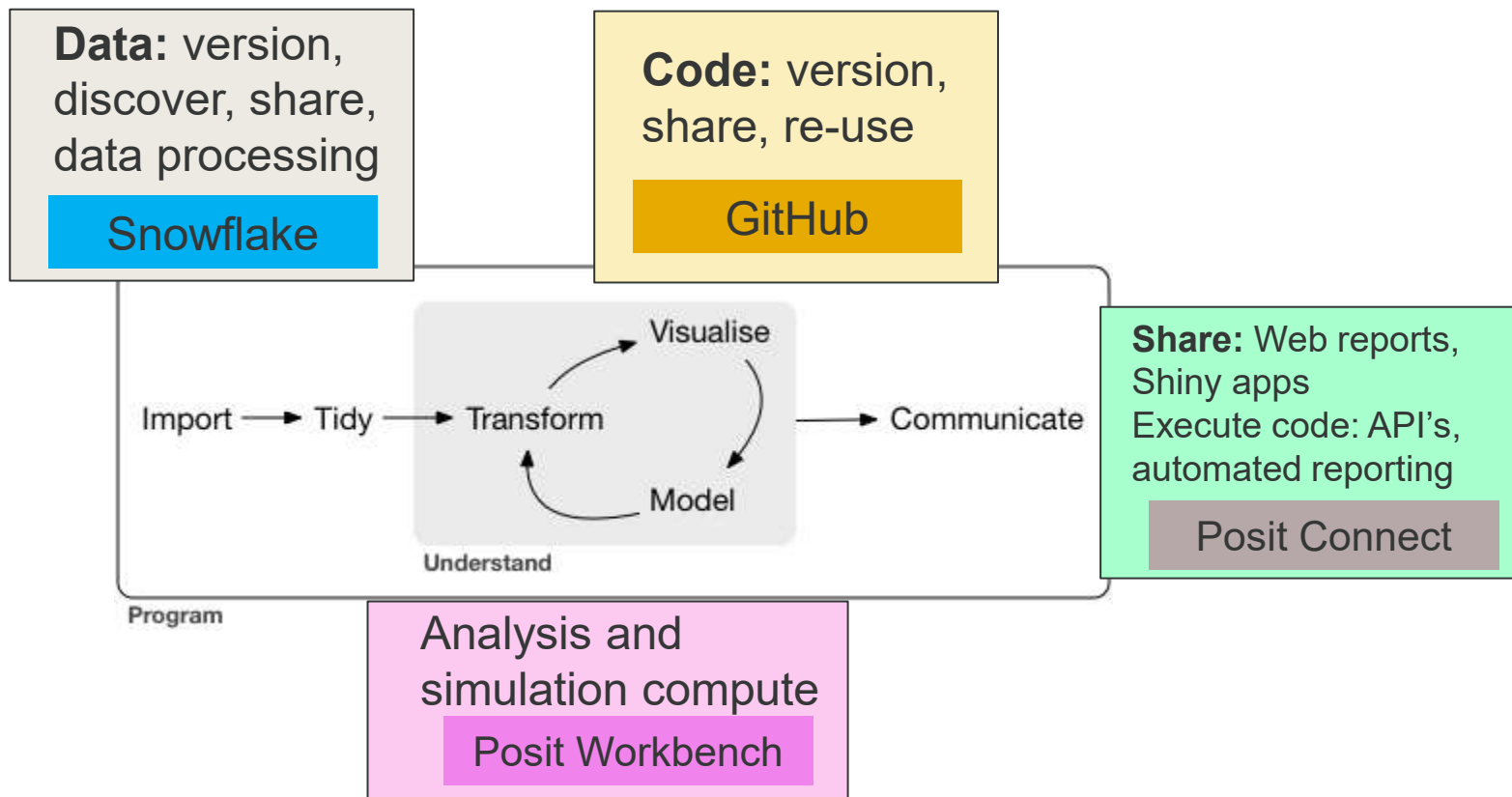
- Pretty seamless
 - GitHub
 - Windows to Linux



Analysis at scale

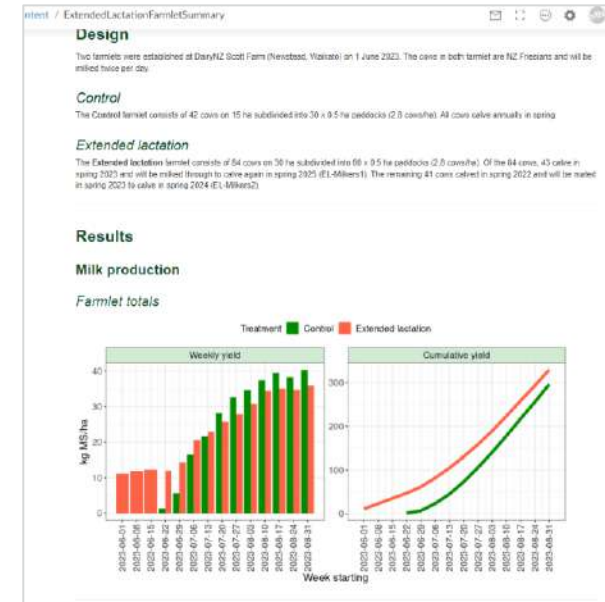
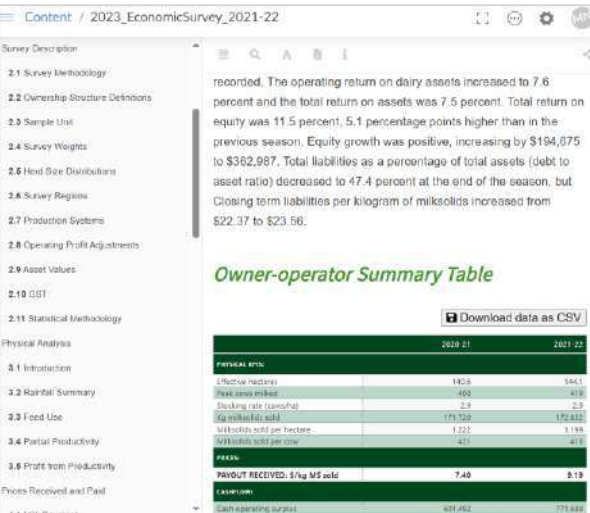


Modern Science Workflow Infrastructure



Posit Connect for distribution

- [Economic Survey](#)
- [Econ Tracker](#)
- [Experiment tracking](#)



Skills & Capability

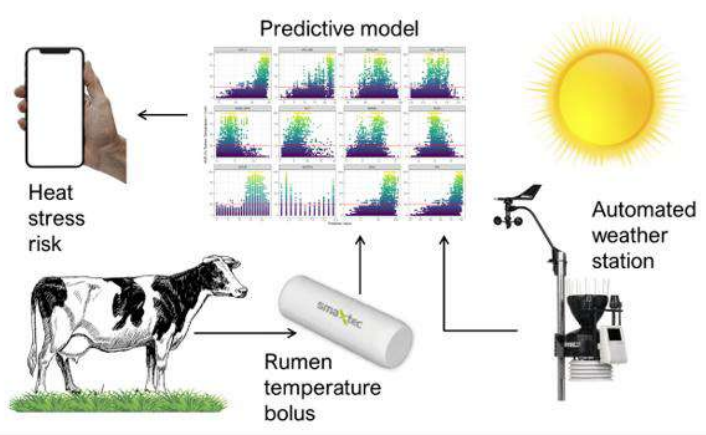
- Statistics Internal course
 - 47 people from 6 teams
- R for Data Science
 - Scientists and Postgrad Students
 - Three times, Waikato Uni, Lincoln Uni
 - Integrate with Learning & Dev. plans
 - Recruiting starts soon



Step 3: Results

- Many examples
 - Data recording for Quality
 - Business continuity
 - ML to predict grazing performance
 - ML to create performance benchmarks for farms

Data recording for quality



Calibration Cuts App



Dairynz Capture Calibration Cut

Calibration

Farm: Scott Farm

Experiment/Treatment: Extended Lactations

Date: 18/08/2024

Save

Paddock	Stage	Quadrat	Ringlit	Visual 1	Visual 2	Visual 3	Visual 4	Weight (g)
ATA	Post	1	5.6	5.75	5.25	5	5.25	24.73
ATA	Post	2	7.3	6.5	6.75	6.5	6.75	34.51
ATA	Post	3	14.4	6.25	6.25	6.5	7.25	38.24
ATA	Post	4	12.9	7.25	7.75	7	8	32.47
AT	Pre	5	24.2	8.5	9.25	9	8.5	47.64
AT	Pre	6	20.1	9.75	10.2	10.5	10.25	69.78
AT	Pre	7	19.9	8.75	8.25	8.5	8.25	42.97
AT	Pre	8	26.3	10.5	10	9.75	10.25	53.84
AT	Pre	9	25.9	12.5	11.25	11	11.75	68.94

Connected Farm

Dairynz

Business continuity

Documenting code

- Previously: loss of key individuals -> Loss of model
- Creating culture where documenting code expected
- Multiple users can now access and run key models

What happens to models when key person leaves?



Before



After

Step 4: Working on and the future

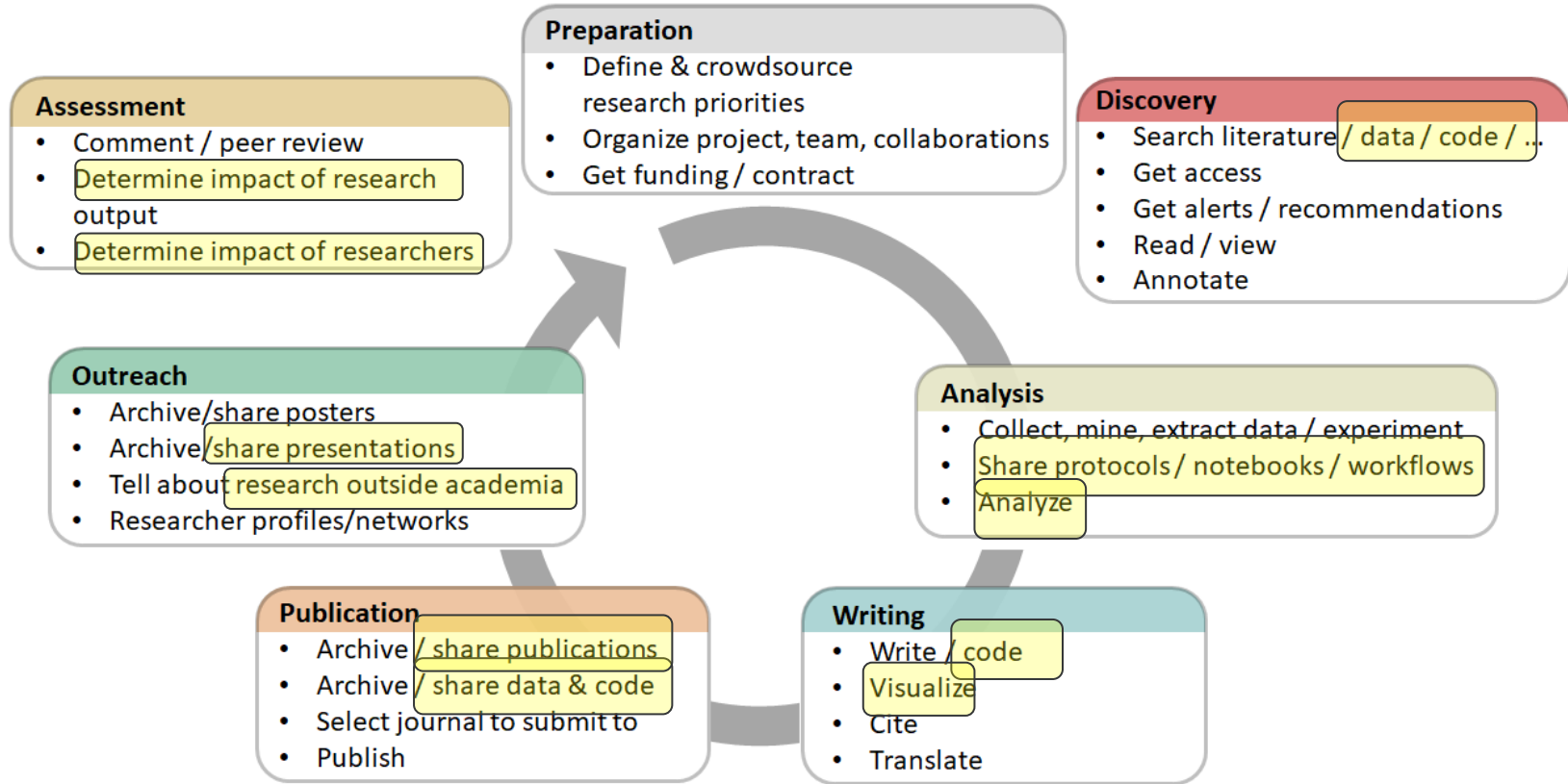
- Continuous improvement
- Open science
- What's next
- Lessons from the revolution

Continuous improvement

- Digital and Research & Science working together
- Celebrate the wins
 - Speed, efficiency, documentation, errors, techniques
- Keep learning
 - R Community, Training
- Link to national and international Initiatives
 - Open Science, Horizon EU
- Use best practice
 - Research and Science goal
 - Digital book – Getting set up for Data Science @ DairyNZ



Modern Science is Open Science



Visit: <https://open-science-training-handbook.gitbook.io/book/open-science-basics/open-concepts-and-principles>

What's next for us?

- From Research to Organisation wide
 - Governance
 - Operating model
 - Prioritisation
 - Change management
- Use cases
 - LLM RAG – researchers, advisers, farmers
 - Paperless data ingestion
 - Metadata quality
 - Copilot
 - Decision support
 - Animal gains from pasture gains

Revolution becomes Business as Usual?

- Getting the foundations right
 - Makes the pain disappear
 - Go a little slower now, to go much faster soon
- Progress doesn't happen by accident!
 - Vision
 - Coalition of the willing
 - Senior buy-in



4th International
**Precision
Dairy Farming
Conference**

3-5 Dec 2025 | New Zealand

precisiondairyfarmingconference.nz



Christchurch



Hosted by *Dairynz*

mark.neal@dairynz.co.nz

