

Jun Huh
Product Manager, NeSI

Mercedes Randell
Research Communities Advisor, NeSI

Collaborating on NeSI's Research Developer Cloud



What is Research Developer Cloud?

Infrastructure-as-a-Service

Enables building custom research tools

Empowers research software engineers



What is IaaS?

“Infrastructure as a service is a cloud computing service model by means of which virtual computing resources are supplied over the Internet.”

Differences to NeSI HPC

- Shared wood fired oven vs access to the whole kitchen
- Queue based vs interactive or long running



Research Developer Cloud

- Infrastructure-as-a-service now accessible to RSEs
- Launched last September with basic building blocks

Our platform's cloud building blocks include:



Compute

Virtual machines optimised for high-performance computational needs. Multiple flavours of CPU and GPU resources to support a range of compute and memory requirements.

→ [Get started with our Compute User Guides](#)



Images

Tailored operating systems to meet your research computing and data needs. Ready-to-use options available, as well as capability to create custom images and contribute to a pool of community-developed images.

→ [Get started with our Images User Guides](#)



Storage

Scalable storage space that can be dynamically mounted to your Compute instances. Options to encrypt storage volumes for added security.

→ [Get started with our Storage User Guides](#)



Networks

Fast, reliable, and secure connectivity built on the [REANNZ national network](#). Options for network customisation and security groups.

→ [Get started with our Networks User Guides](#)



Identity

Identity management services to create application credentials and control access to your projects.

→ [Get started with our Identity User Guides](#)



Application Programming Interface (API)

All services are programmable via a public API to enable repeatable definition of infrastructure through software code.

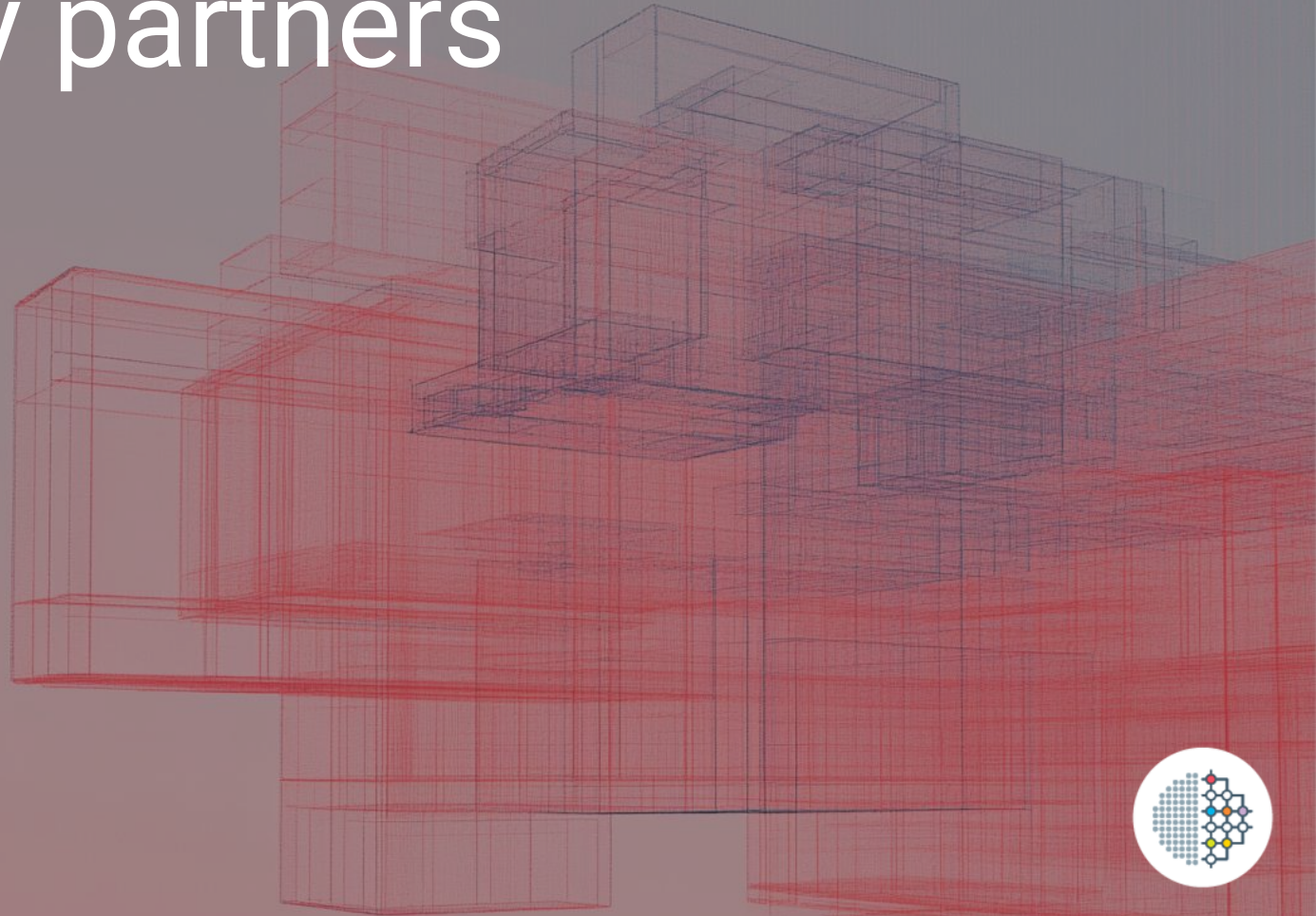


Our focus / differentiators

- Focusing on collaboration, building networks, and community empowerment
- Focus on good developer experience
- Building shared knowledge and expertise → Blueprints and patterns
- Adjacency to existing NeSI services



Our early partners



Tenancy updates

- 3 new tenants
 - AUGMENT project from Auckland Bioengineering Institute
 - GNS Te Pū Ao (Institute of Geological and Nuclear Sciences)
 - Manaaki Whenua – Landcare Research
- Lots of internal projects
 - NeSI training environment
 - NeSI migration



AUGMENT project

- A group from Auckland Bioengineering Institute
- Research context: Machine learning based video analysis toolchain for early detection of neurological disorders
- Cloud use case: Back-end server for hosting video data
- Edition, a mobile development team is building a mobile app
- Mixed use of HPC and cloud
 - Cloud back-end for the mobile app to receive and store video data
 - Processing on NeSI HPC to build ML models



Christof Mueller's group from GNS

- Research context: AI and advanced analytics team working on various research projects such as Rapid Characterisation of Earthquakes & Tsunami (RCET)
- Cloud use case: Deploying Kubeflow, and applications that can be used by Emily Lane, researcher from NIWA
- Ciaran King, an RSE from GNS, has deployed Kubeflow, and have set up GNS Portainer to talk to NeSI Cloud VM
- Interests around GPU access and IdM for research applications accessible by wider GNS researchers



Manaaki Whenua Landcare Research

Project 1

- Research context: LUMASS (the Land-Use Management Support System)
- Cloud use case: Set up a pipeline for building models that feeds into LUMASS
- Consultancy project where Chris and Kahu from NeSI worked together with Alex Herzig and Chris Wydra from MWLR
- Nick Spencer is looking for more cultural and paradigmatic shift in the organisation, rather than just building services to be used.



Manaaki Whenua
Landcare Research



Manaaki Whenua Landcare Research

Project 2

- Research context: DeepSeg project around remote sensing
- Cloud use case: Deploying the pipeline into the cloud and improving to make it more useable by others
- Challenges: Passing workload into NeSI HPC
- Team working actively to build proof of concept



Manaaki Whenua
Landcare Research



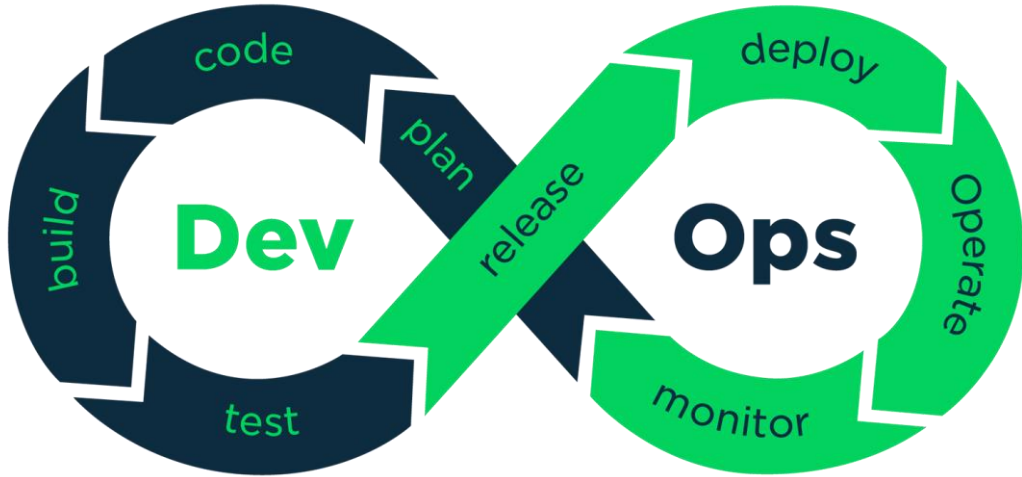
What is IaC?

“Infrastructure as code is the process of managing and provisioning computer data center resources through machine-readable definition files, rather than physical hardware configuration or interactive configuration tools.” - Wikipedia

- Just cooking vs writing a recipe
- Transparency
- Repeatable
- Enables collaboration



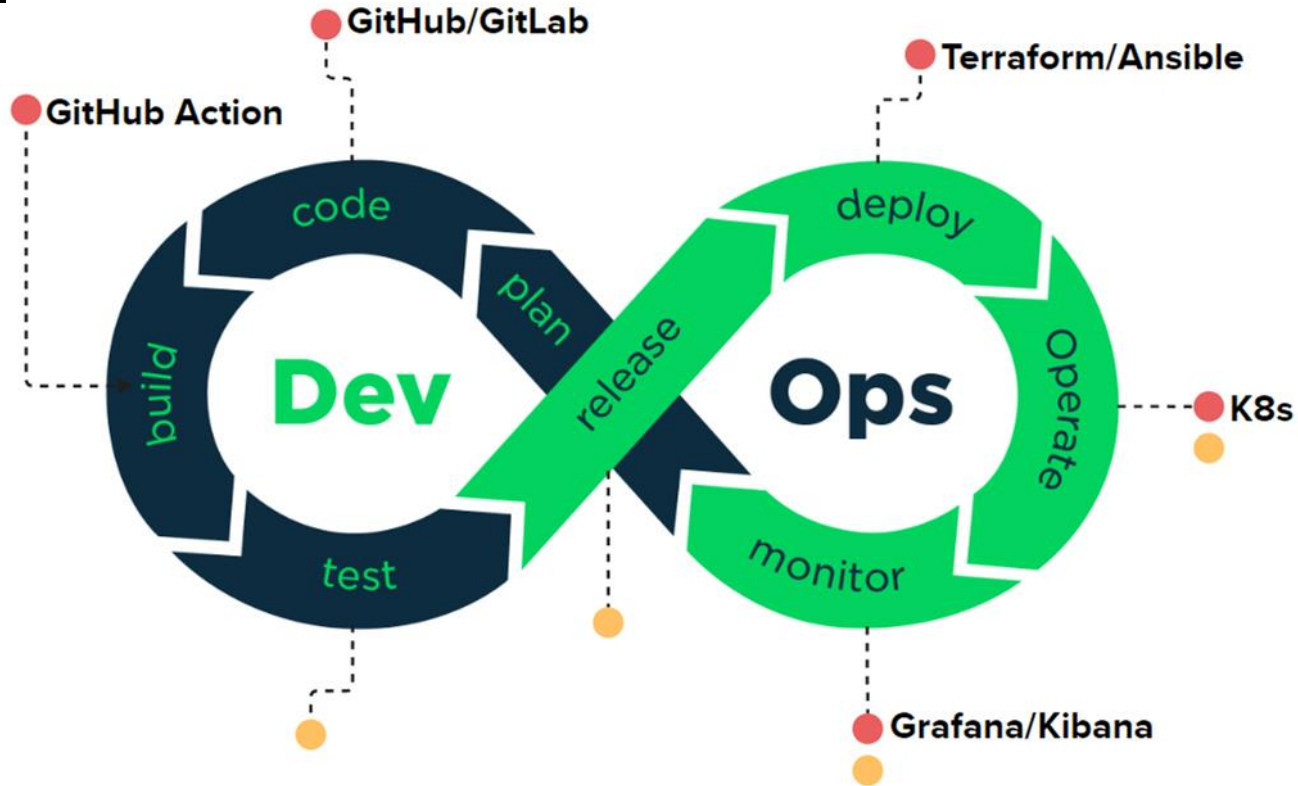
DevOps Workflow



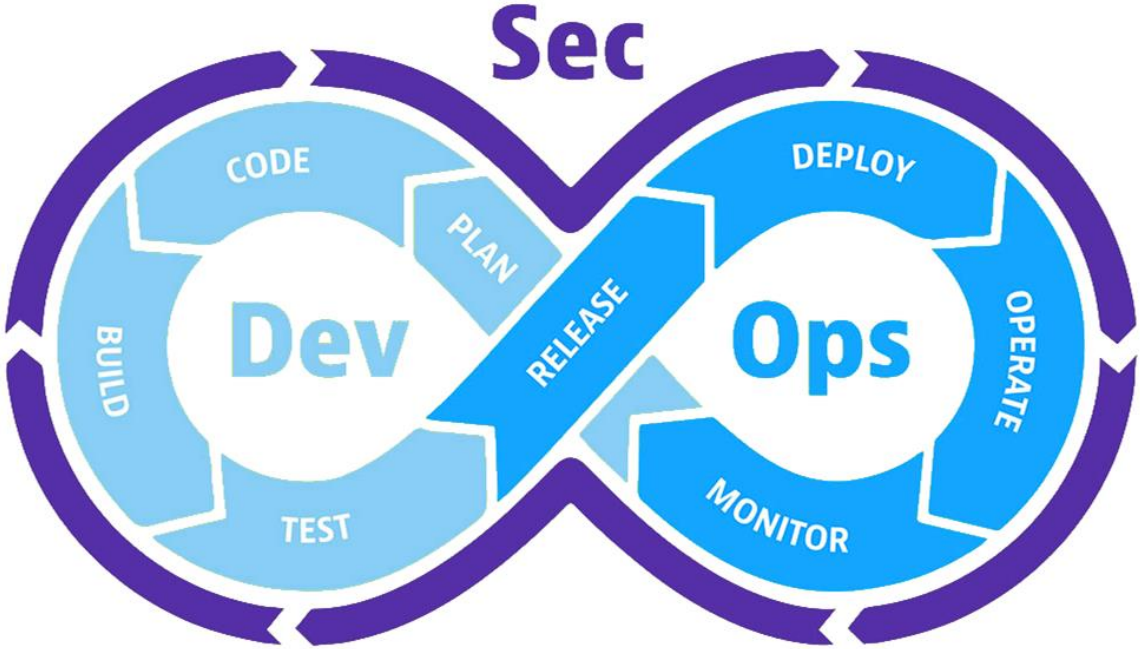
The DevOps workflow. Source: Pease, 2017.



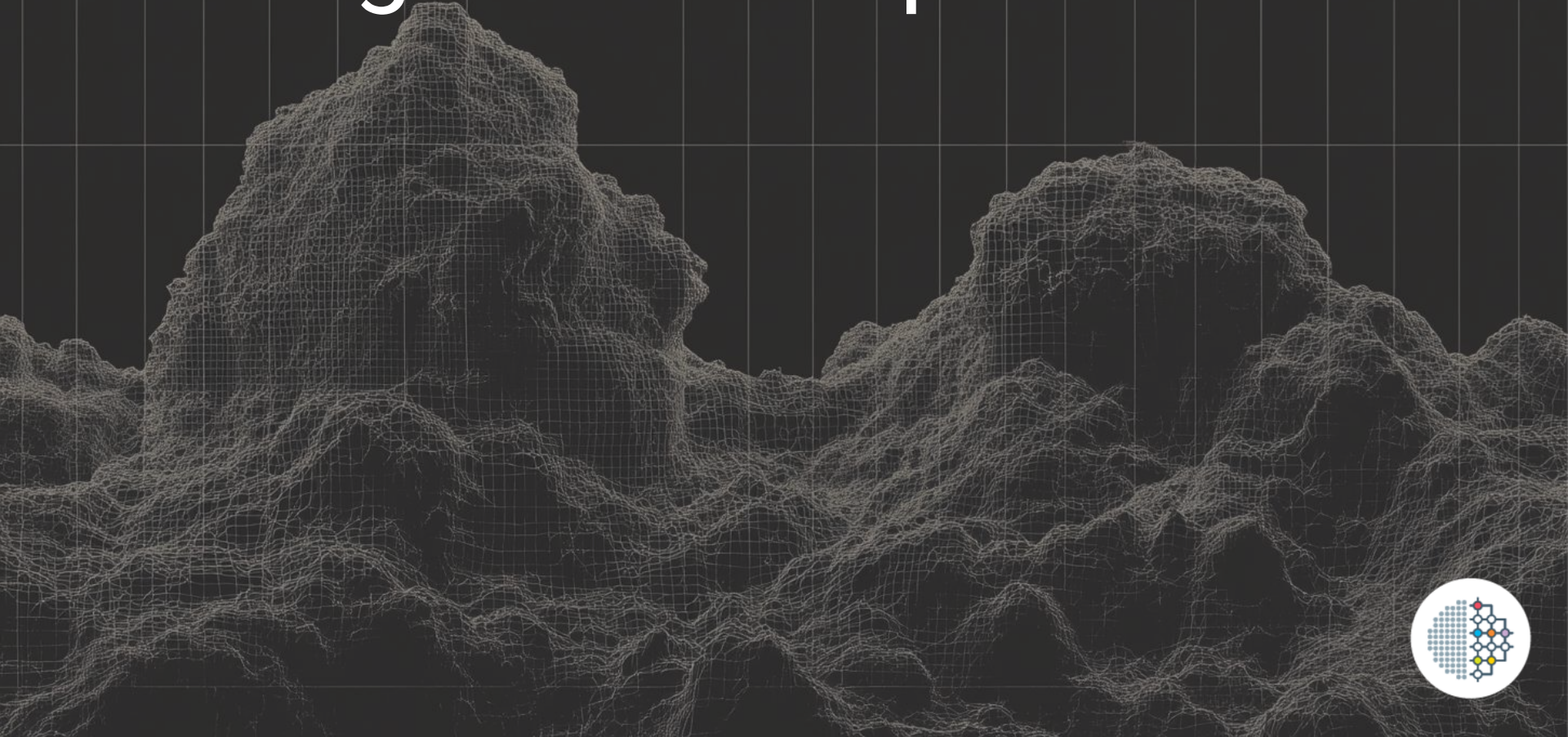
DevOps Workflow



DevSecOps



Looking into example scenarios



AGDR (Aotearoa Genomic Data Repository)

- Research context: Collaboration with Genomics Aotearoa. Building a long term data repository of genome sequencing data for taonga species
- Challenges: Data and platform onshore, kaitiakitanga, FAIR
- Technology used: Gen3 (Open source genomic domain data repository solution, University of Chicago), Kubernetes (for platform management, health and scaling), Google Forms / Zendesk (for application process)



genomics
aotearoa

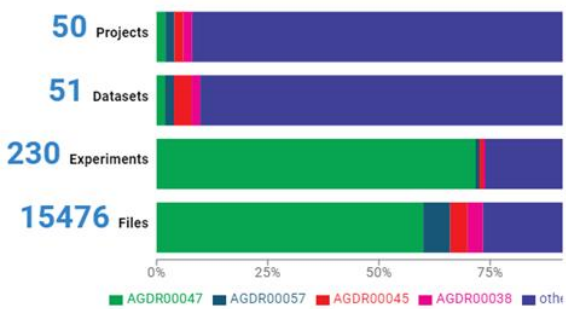




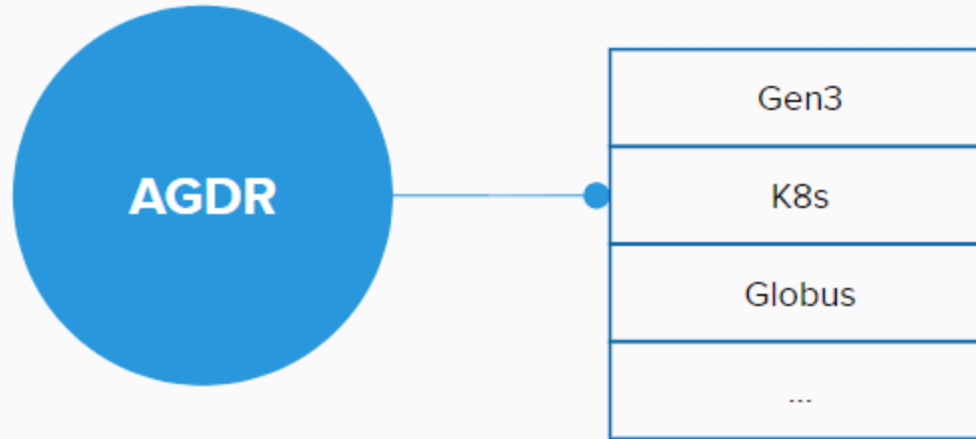
kākāpō chick photo by Dianne Mason, 2009 CC2.0

Aotearoa Genomic Data Repository

The Aotearoa Genomic Data Repository provides secure within-nation storage, management and sharing of non-human genomic data generated from biological and environmental samples originating in Aotearoa New Zealand. This resource has been developed to follow the principles of Māori Data Sovereignty, and to enable kaitiakitanga (guardianship), so that iwi, hapū and whānau (tribes, kinship groups and families) can effectively exercise their responsibilities as guardians over biological entities that are taonga (precious or treasured). While the repository is designed to facilitate the sharing of data – making it findable by researchers and interoperable with data held in other genomic



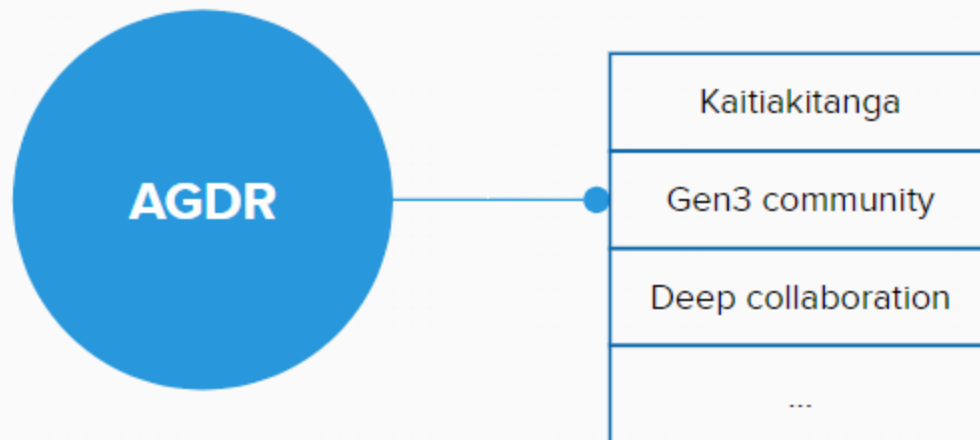
Building shared knowledge through collaboration



genomics
aotearoa



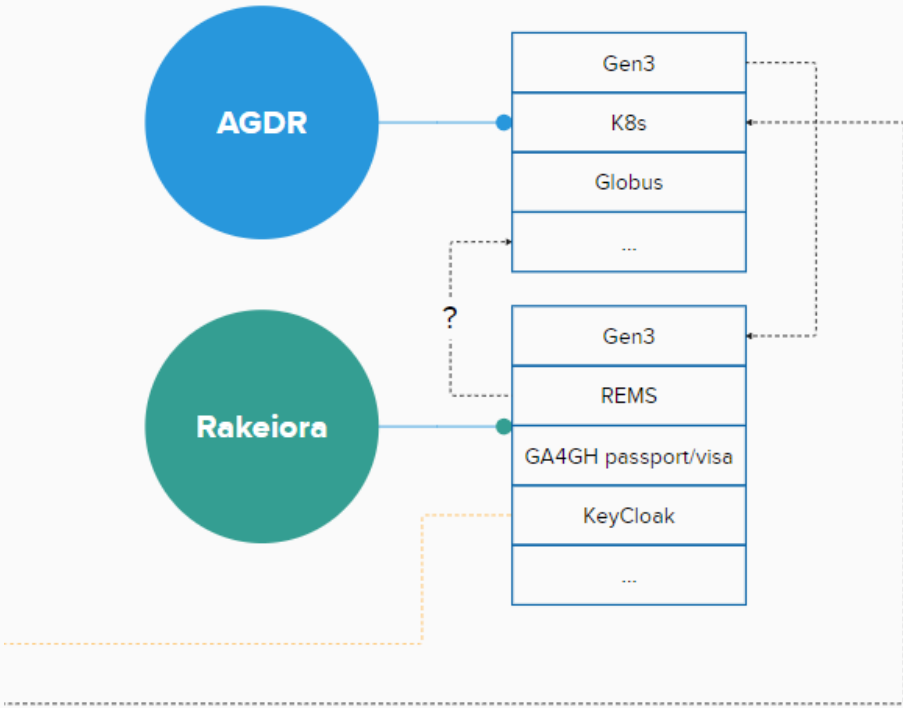
Building shared knowledge through collaboration



genomics
aotearoa



Building shared knowledge through collaboration



Manaaki Whenua Landcare Research

- Research context: LUMASS (the Land-Use Management Support System)
- Cloud use case: Set up a pipeline for building models that feeds into LUMASS
- Consultancy project where Chris and Kahu from NeSI worked together with Alex Herzig and Chris Wydra from MWLR
- Nick Spencer is looking for more cultural and paradigmatic shift in the organisation, rather than just building services to be used.



Manaaki Whenua
Landcare Research



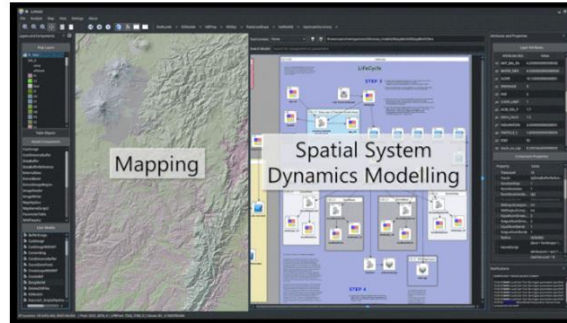
Manaaki Whenua Landcare Research



📍 New Zealand
🌐 Manaaki Whenua
🔄 Repository



LUMASS - Land-Use Management Support System



LUMASS is designed to provide support for two high level aspects of land management: i) land use impact assessment and ii) spatial planning. The former aspect is supported by its [spatial system dynamics modelling framework](#) and the latter by its [spatial optimisation framework](#).



Manaaki Whenua
Landcare Research



Manaaki Whenua Landcare Research

- Main outcomes
 - Built familiarity with NeSI's cloud platform and related infrastructure as code tools among Manaaki Whenua research software engineers
 - Developed a proof of concept showing how GitHub Actions and NeSI's research cloud platform could be used to automate the LUMASS build process
 - Created a blueprint showing how a similar approach could be used for other projects
 - <https://github.com/chrisdjscott/flexi-actions-runner-demo>



Manaaki Whenua
Landcare Research



Manaaki Whenua Landcare Research

- Main outcomes
 - Built familiarity with NeSI's cloud platform and related infrastructure as code tools among Manaaki Whenua research software engineers
 - Developed a proof of concept showing how GitHub Actions and NeSI's research cloud platform could be used to automate the LUMASS build process
 - Created a blueprint showing how a similar approach could be used for other projects
 - <https://github.com/chrisdjscott/flexi-actions-runner-demo>



Manaaki Whenua
Landcare Research



Manaaki Whenua Landcare Research

"The automated build process promises to greatly simplify the production of new releases including artifact management. It will provide a more stable build environment and reduce the reliance on my personal laptop for building LUMASS. Last, not least, It saves valuable time I can use for science."

Alex Herzig



"Manaaki Whenua have been pleased to have had the opportunity to begin understanding how these types of remote platforms can broaden the type of compute we can offer our researchers. This proof of concept provided the opportunity to learn more about researcher needs when migrating to these types of compute platform."

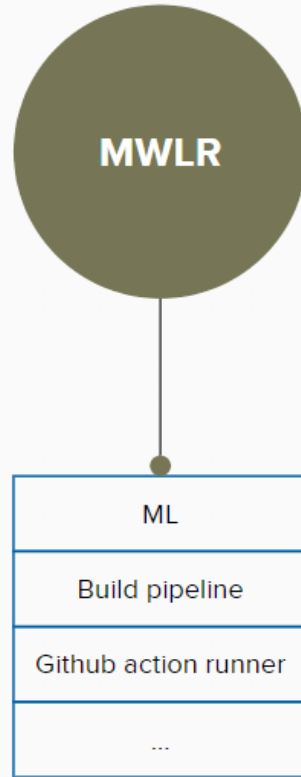
Nick Spencer



Manaaki Whenua
Landcare Research



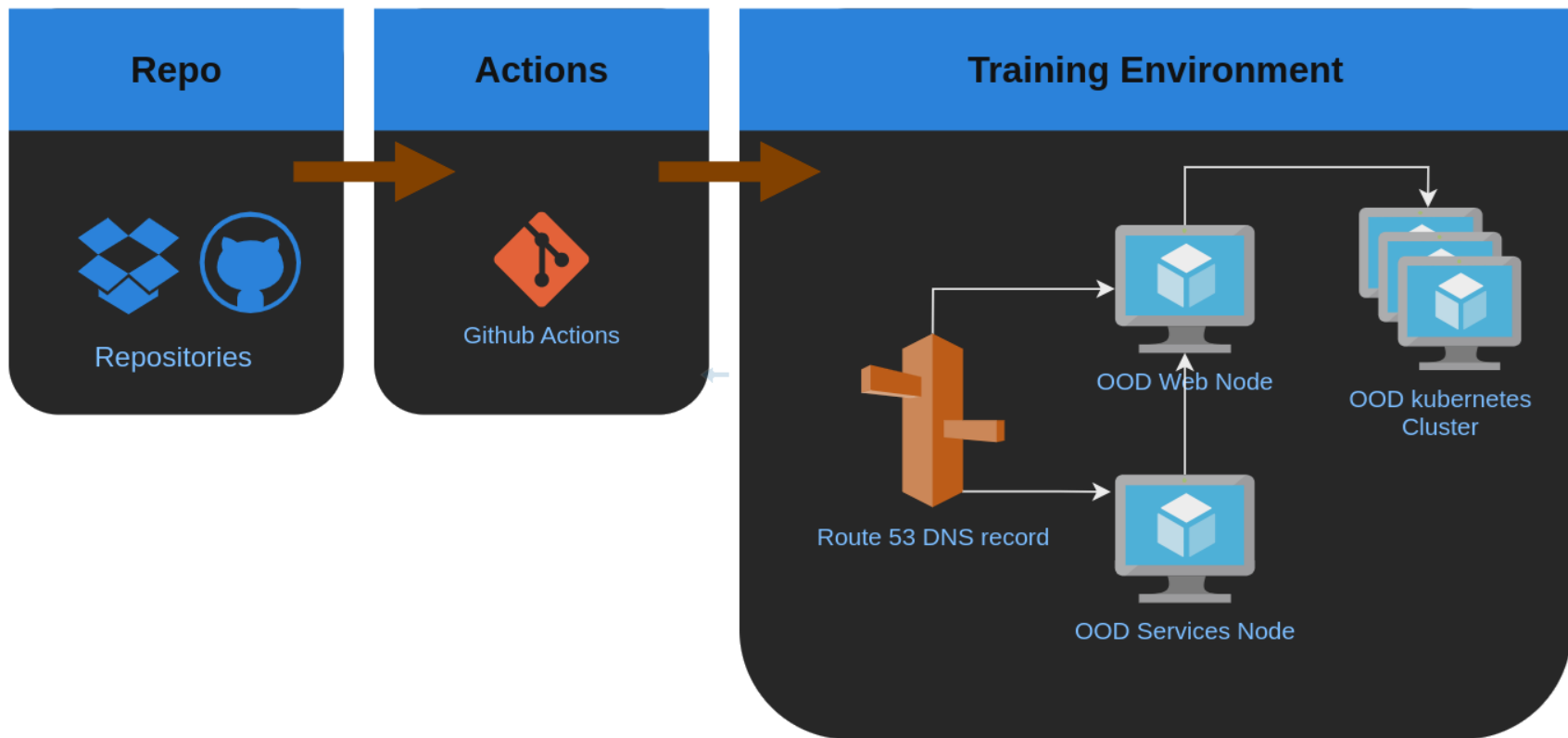
Building shared knowledge



Internal project: NeSI's training environment

- Started off as a learning experience for NeSI
 - Applying repeatable code using Ansible and Terraform
 - A way to provide training outside of NeSI
- Delivered over 20 training events using the new training environment
- Technology used
 - Open OnDemand, Keycloak
- Tooling
 - Terraform, Ansible, Github action
 - launch, use, and destroy pattern






NeSI training environment

- Training Environment Git Repo
 - <https://github.com/nesi/training-environment>




NeSI Training Environment


Training Course Apps




ML101
System Installed
App




ML102
System Installed
App




RNA-Seq
System Installed
App




scRNA-Seq
System Installed
App



RStudio
System Installed
App




**Intermediate shell
for bioinformatics**
System Installed
App




**Containers
workshop**
System Installed
App




**Introduction to
Nextflow Workshop**
System Installed
App




code-server app
System Installed
App



Python
System Installed
App



**Introduction to
Shell (for
bioinformatics)**
System Installed
App



RStudio-IndigiData
System Installed
App

Active interactive sessions [view all \(1\)](#)

RNA-Seq (rnaseq-tjv6dmb) **2 cores** | Running

Host: 10.1.0.47

✖ Cancel

Created at: 2024-10-16 09:03:47 NZDT

Time Used: 4 hours and 50 minutes

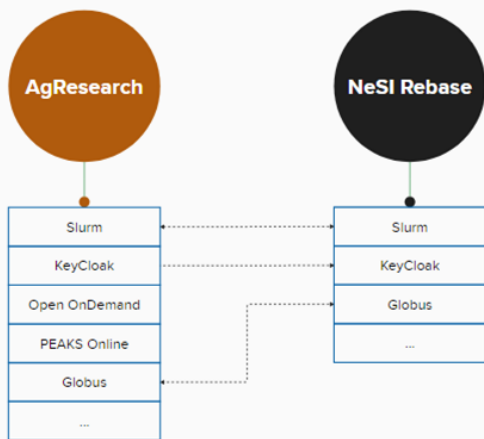
Session ID: 1ee3307f-fbb3-4b2a-bcaf-55d251be4015

🔗 Connect to RNA-Seq RStudio

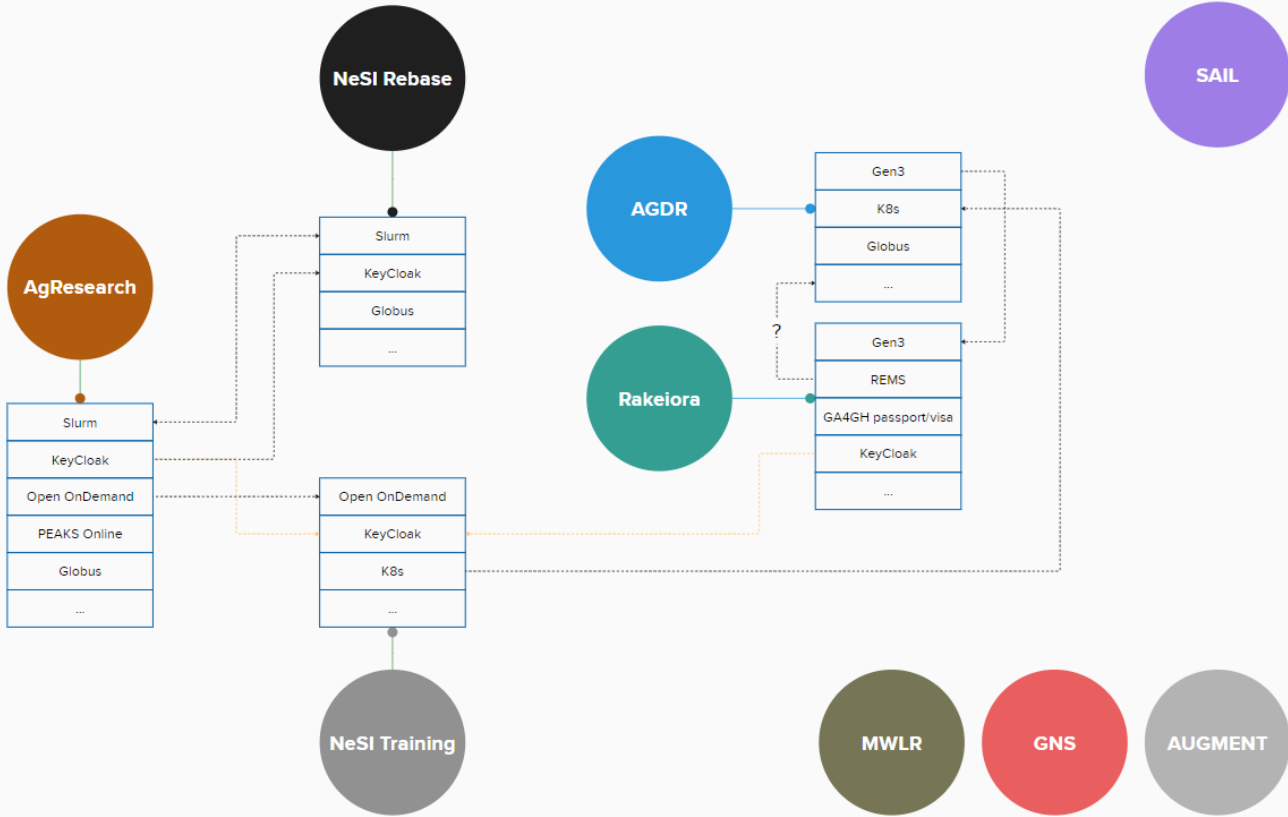


NeSI migration

- NeSI is in the process of building its next HPC
- Collaboration across teams
- Learnings that can be shared back
 - Services and tools that are being built as well as those used to help building



Network of knowledge



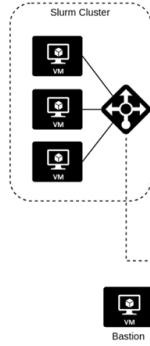
NeSI migration: NeSI OnDemand

- Replacing NeSI's JupyterHub environment for interactive applications
- Open OnDemand, Keycloak, NFS, FreeIPA
- Tooling
 - Terraform, Ansible, GitLab pipelines, K8s, Secret management with Hashicorp Vault

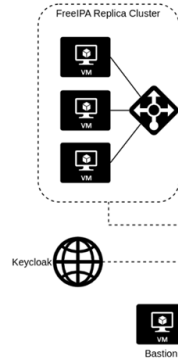


nesi-rcd

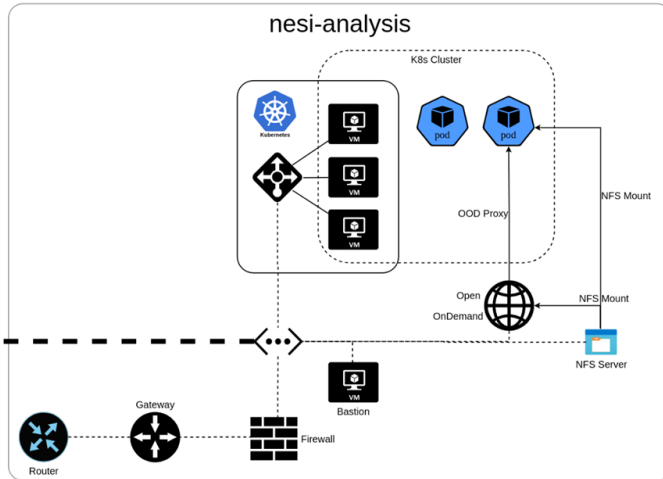
nesi-slurm










nesi-freeipa



nesi-analysis




























































Search or go to...

Project

- ood-analysis-environment
- Pinned
- Issues 0
- Merge requests 0
- Manage
- Plan
- Code
- Build
- Pipelines**
- Jobs
- Pipeline editor
- Pipeline schedules
- Artifacts
- Secure
- Deploy
- Operate
- Monitor
- Analyze
- Settings
- Help

 00:00:52 4 days ago	#1451578553  main → 24d2f298  latest			
 00:01:05 4 days ago	Merge branch 'jupyter-v0.9.0' into 'main' #1451577494  main → 24d2f298  latest			
 00:00:50 5 days ago	update Jupyter to version 0.8.0 #1449850755  main → d8f7df37 			
 00:01:00 5 days ago	update Jupyter to version 0.8.0 #1449847626  main → d8f7df37 			
 00:00:58 6 days ago	Merge branch 'jupyter-v0.7.0' into 'main' #1448146554  main → a28ce6cf 			
 00:01:02 6 days ago	Merge branch 'jupyter-v0.7.0' into 'main' #1448139921  main → a28ce6cf 			
 00:00:53 1 week ago	Merge branch 'HPC-188-bring-central-monitor...' #1446790466  main → eb5a519c 			
 00:00:56 1 week ago	Merge branch 'feature/update-jupyter-app-0.6...' #1446789666  main → 8816da4c 			
 00:00:57 1 week ago	Merge branch 'feature/update-jupyter-app-v0...' #1446781604  main → b6de7a9a 			
 00:01:00 1 week ago	Merge branch 'feature/update-jupyter-app-v0...' #1446775557  main → b6de7a9a 			

Navigation icons: Home, Search, GitLab logo, and a search bar with the text "Search or go to..."

- Project
 - ood-analysis-environment
 - Pinned
 - Issues (0)
 - Merge requests (0)
 - Manage
 - Plan
 - Code
 - Build
 - Secure
 - Deploy
 - Operate
- Environments
 - Kubernetes clusters
 - Terraform states
 - Terraform modules
 - Google Cloud
 - Monitor
 - Analyze
 - Settings
- Help

Environment management interface for 'ood-analysis-environment'.

dev [Open] [Stop]

Success Latest Deployed #138 d27e4638 Deployed 4 weeks ago

Updates to tell which env you are in

Triggerer	Job	Branch
@lbrick	configuration-dev	main

Waiting #139 d27e4638 Created 4 weeks ago

Updates to tell which env you are in

Triggerer	Job	Branch
@lbrick	deploy_...l_apps-dev	main

production [Open] [Stop]

sandbox [Stop]

test [Open] [Stop]

Success Latest Deployed #208 24d2f290 Deployed 4 days ago

Merge branch 'jupyter-v0.9.0' into 'main'

Triggerer	Job	Branch
@lbrick	deploy-...vironment	main

Waiting #140 d27e4638 Created 4 weeks ago

Updates to tell which env you are in

Triggerer	Job	Branch
@lbrick	provision-test	main

What's next?



Sharing the knowledge

Blueprints and patterns

- Kubernetes
- *Globus*
- *Instrument pipeline*
- *Open OnDemand*
- ...

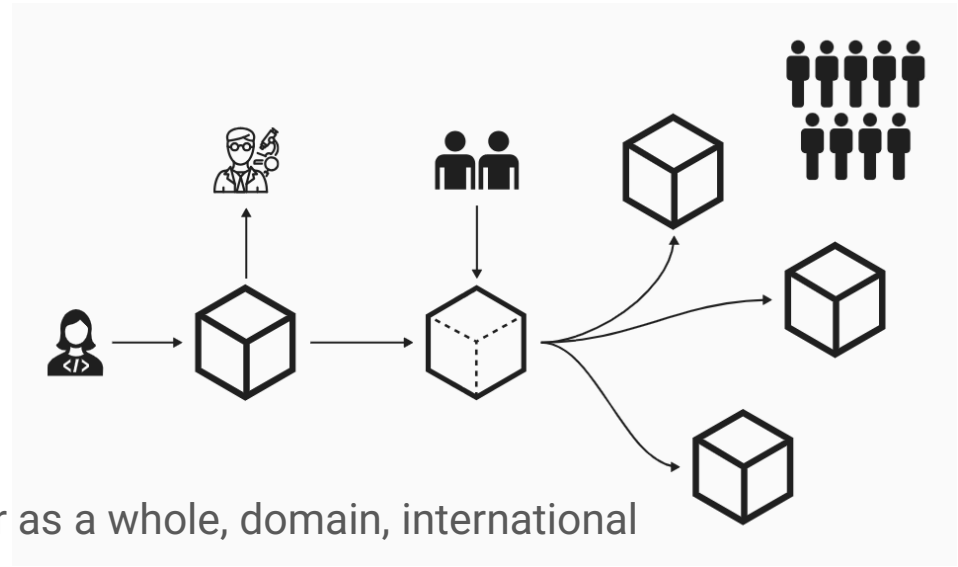
Researchers would have access to shared solutions and common tool chains for high value common use cases, e.g. notebooks, VLS, national data transfer platform etc.

Ecosystem for RSE/DevOps to build solutions and easily share not just with their own institutional researchers but across the sector



Pathway for RSEs

- Pathway for RSEs
 - Contribution to the group
 - To the organisation
 - Wider communities, the sector as a whole, domain, international
- How?
 - DevOps, open standards, community building tools
 - GitLab, Terraform/Ansible, K8s...
 - Learning from others; ARCOS, etc.





[Home](#)

Research Developer Cloud

[Table of contents](#)

[Get started](#)



Cloud-native services

Use NeSI's cloud infrastructure and on-demand services to develop and manage custom, interactive solutions.



Programmable infrastructure

Apply DevOps practices enabled by Infrastructure as Code (IaC) to automate, measure, collaborate, and learn.



Partnership-led approaches

Partner with NeSI's DevOps specialists to build a platform or tools that can benefit your research community.

NeSI is building a research cloud computing platform for easier and more adaptable collaboration around research data.

Research teams can use our Research Developer Cloud to develop novel solutions that enable research.

Get started

[Scroll down to see how to apply for early access](#)

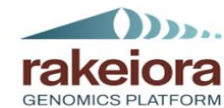
Talk to us

[If you want to chat with us about how our platform can support your research](#)

[Get started](#)



REANVZ



StackHPC

