UWA Document (Public)

APeX SUMMIT

BUILDING THE GUARDRAILS

How we use Globus Flows and Globus Compute to enforce access control and data hygiene

Jack Bryant | Solutions Architect, The University of Western Australia

THE UNIVERSITY OF WESTERN AUSTRALIA

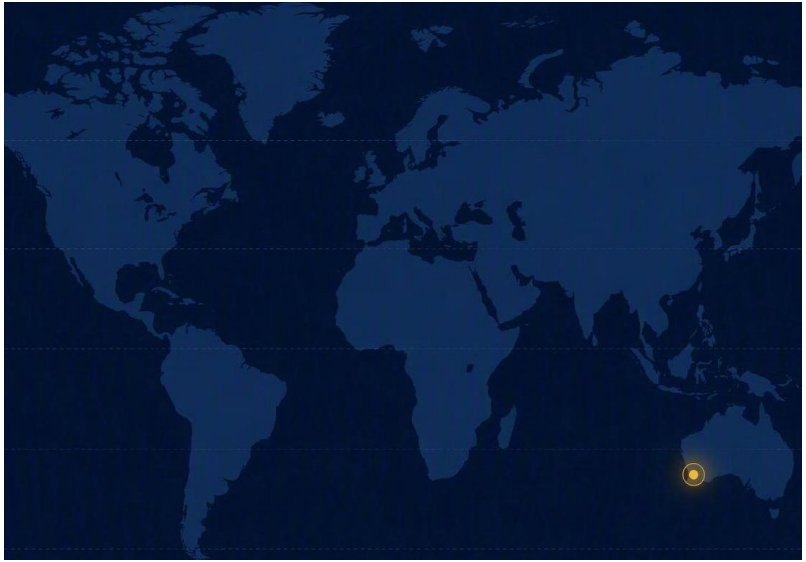
SEEK WISDOM

This is a UWA Official document authorised or intended for public ac

FROM ABOUT AS FAR AWAY AS POSSIBLE

3,300 km

west of Sydney — and a long way from most of this room. When people think of Australia, they think Sydney. We are muuuuuuch further west.



The University of Western Australia, Perth - WA

THE LANDSCAPE

Between our shared instruments and our HPC centre, two gaps sat wide open.

SHARED DATA STORES

Instrument data is copied to a shared store. No tamper controls, no separation of privilege between researchers — and we have seen incidents of theft.

NO DATA HYGIENE

No hygiene engine sat anywhere in the path to our HPC centre. The threat isn't only tampering or theft — it's malicious code, RATs, reverse shells.

THE AI ERA

Creating and proliferating those tools has never been cheaper or easier. The volume of unchecked content arriving only rises from here.

Both are controls that have to happen somewhere in the data's journey.

ACCESS CONTROL, DERIVED FROM THE DATA

Instrument to Mapped Collection to Guest Collection — with the ACL built from the dataset itself.

1 PREFIX

The source dataset on the instrument is prefixed with the owner's unique staff or student number.

2 BACK UP

A Globus Flow automatically backs up all instrument data to a Mapped Collection.

3 PARSE

The same Flow enumerates the folder name to extract the staff or student ID.

4 RESOLVE

A Globus Compute function converts that University ID into the owner's Globus identity.

5 SHARE

A Guest Collection or ACL is created, sharing the dataset directly with that Globus ID.

Nobody files a request. Ownership is carried by the data, and the Flow does the rest.

UWA Document (Public)

THE HYGIENE ENGINE

How a Globus Compute function became the data hygiene engine in front of HPC.

Staging, not destination

Every HPC collection is fronted by a Flow. Inbound data lands in a staging path, never the final destination.

One run, one identity

The Flow mints a unique run ID, so a dataset is traceable from arrival through to release.

Compute does the work

The Flow finds a live Globus Compute endpoint on the scan host and submits the job next to the data.

Defender does the inspection

The registered function expands archives and runs Microsoft Defender across the staged content.

Hygiene, not rejection

Defender remediates what it finds in staging, so what moves on is already clean. No transfer is bounced.

Runs as the researcher

The endpoint runs multi-user, so the scan executes as the HPC user. No privilege escalation, ownership intact.

This is a UWA Official document authorised or intended for public access.

THE PATTERN

Two different problems, one shape. Both controls are built from the same three moving parts.

YESTERDAY

Network controls

We solved these problems at the perimeter, with a firewall between us and the risk.

→

TODAY

Identity controls

Federated identity, MFA and audit now carry most of that weight.

→

NEXT

Workflow controls

A trigger, a Globus Compute function that makes the decision, and an auditable run record.

Compliance becomes automation. Governance becomes automation. And researchers never have to think about it.

UWA Document (Public)

THE CONCLUSION

Cyber tends to see network threat vectors. In practice the picture is the other way around.

TRANSFER — **SCAN** — **APPROVE** — **RETAIN** — **DISCOVER**

Globus is the most trusted and verifiable part of our research data and compute landscape — so it belongs inside our compliance and cyber story, not outside it.

Thank you.

This is a UWA Official document authorised or intended for public access.