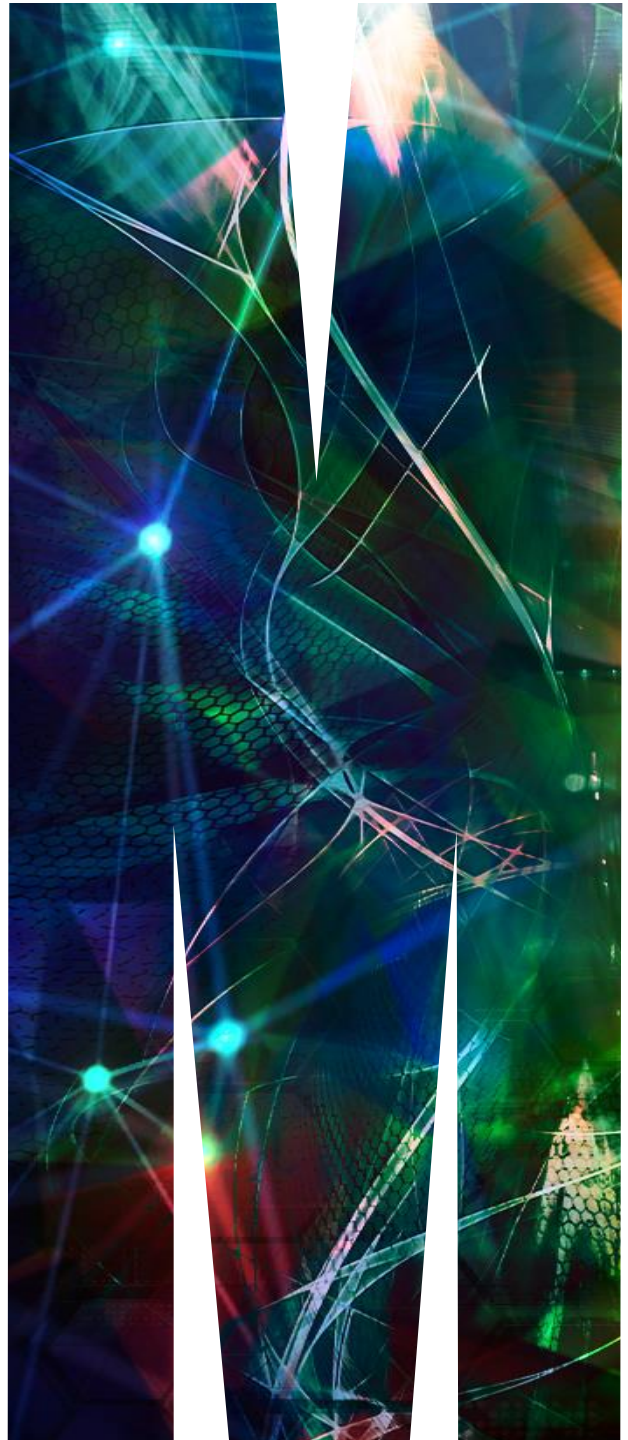
MONASH University

Eyes off! Validating research on illegal data through the Data Airlock.

Dr Gregory Rolan

Thursday 19 October

eResearch Australasia, 2023 Conference, Brisbane Convention & Exhibition Centre



The AiLECS Lab

Artificial Intelligence for Law Enforcement and Community Safety

- Multi-disciplinary research centre collaboration between AFP and Monash
- Founded in 2019; primary motivator - technological countering of online child exploitation
- Research Themes:
  - AI and machine learning for community safety
  - Ethics, explainability and transparency of AI
  - AI in forensics
  - AI and the legal system.
  - Pattern of life characterisation
  - Data ontologies
- Research infrastructure & process

Problem domain

Sharing restricted data

- Machine learning requires collaboration between wide variety of stakeholders
- Dynamic sensitivities in data (& models)
  - Regulation
  - Policy
  - Community expectations
- Not a new problem; predates AI paradigm
- Range of social and technical measures

Dealing with sensitive data

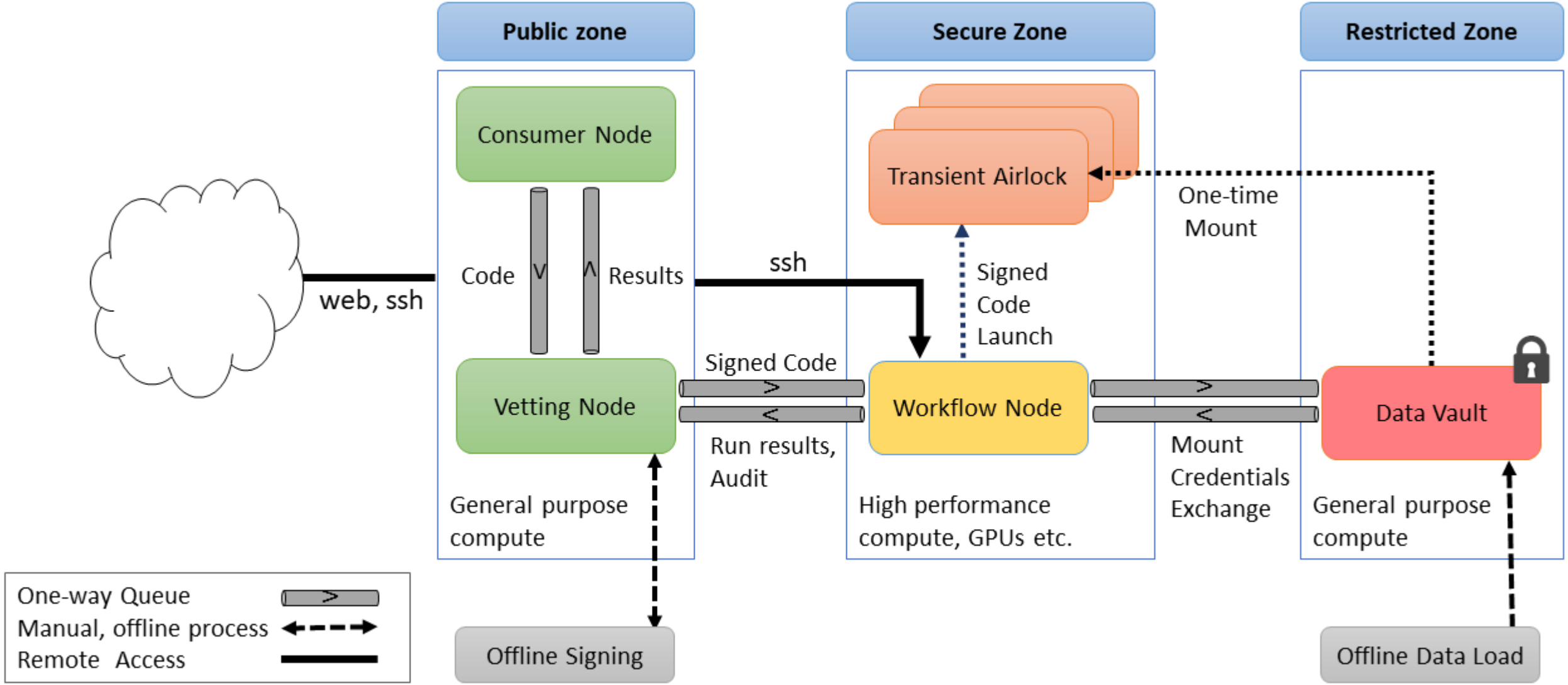
- Social measures
- Anonymisation methods
- Provably Secure Protocols (inc. Homomorphic encryption)
- Trusted Execution Environments
- Secure analytics platforms

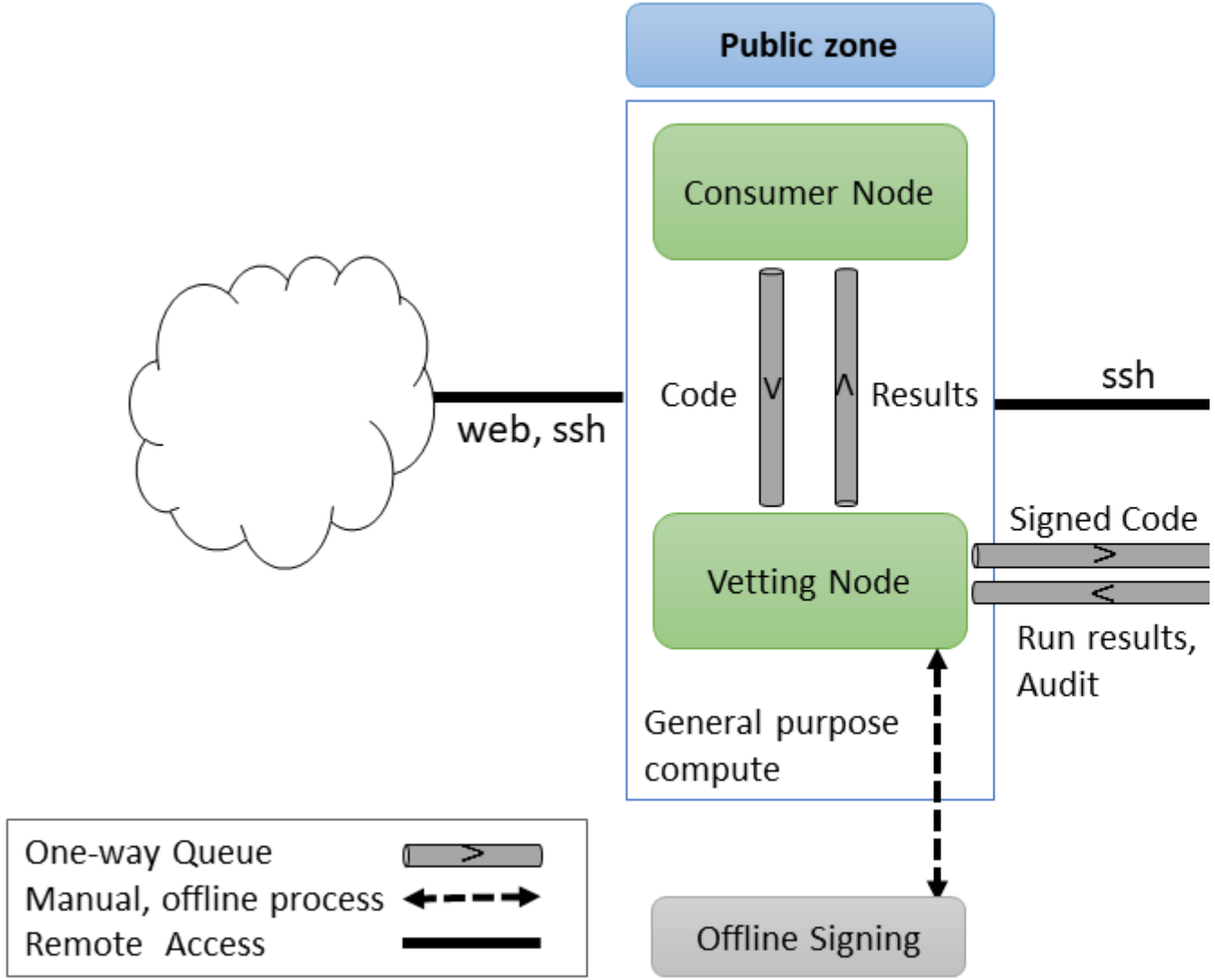
Assumption: Data can be shared in some form

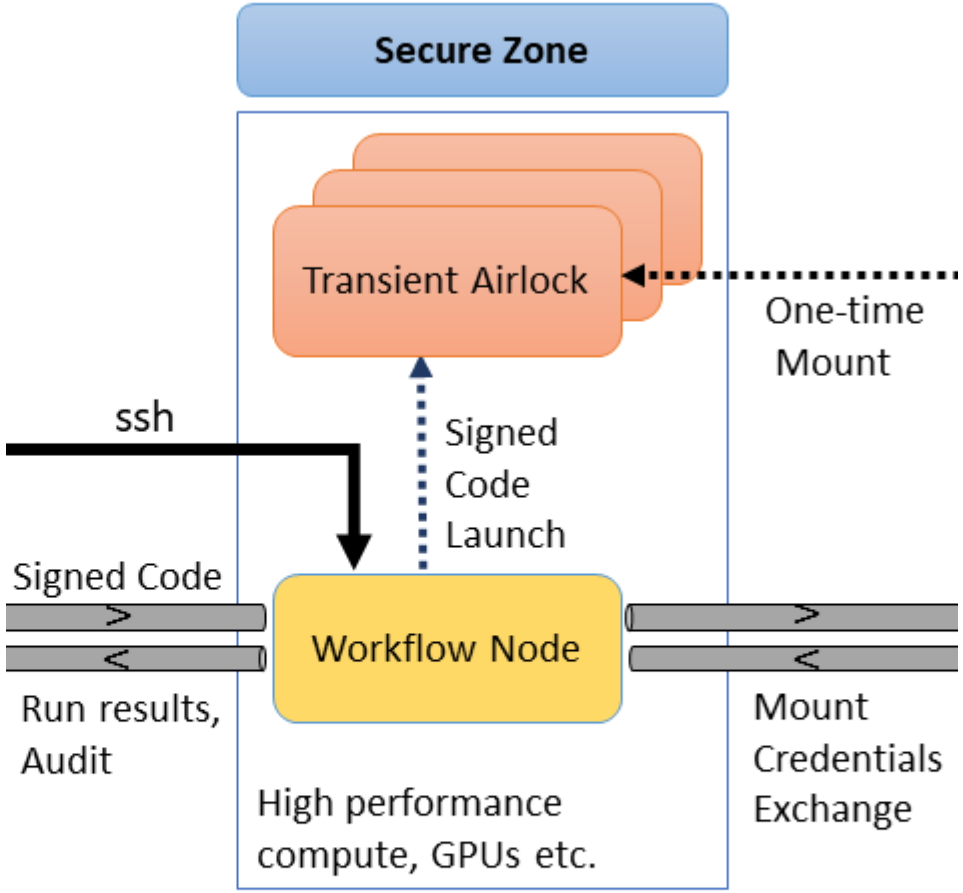
The CSAM use-case

Significant, trans-national problem

- Overwhelming volume of cases / media items
  - Structured data; Text; Audio; Images; Video; Binary; Contextual metadata
- Secondary trauma
- Need for automated triage (machine learning classification)
- Possession of CSAM is proscribed by Commonwealth; State law
  - Research use is specifically not a defence.


Public zone
Secure Zone
Restricted Zone
Consumer Node
Transient Airlock
One-time
Mount
web, ssh
Code
Results
ssh
Signed
Code
Launch
Signed Code
Vetting Node
Workflow Node
Data Vault
Run results,
Audit
Mount
Credentials
Exchange
General purpose
compute
High performance
compute, GPUs etc.
General purpose
compute
One-way Queue
Manual, offline process
Remote Access
Offline Signing
Offline Data Load

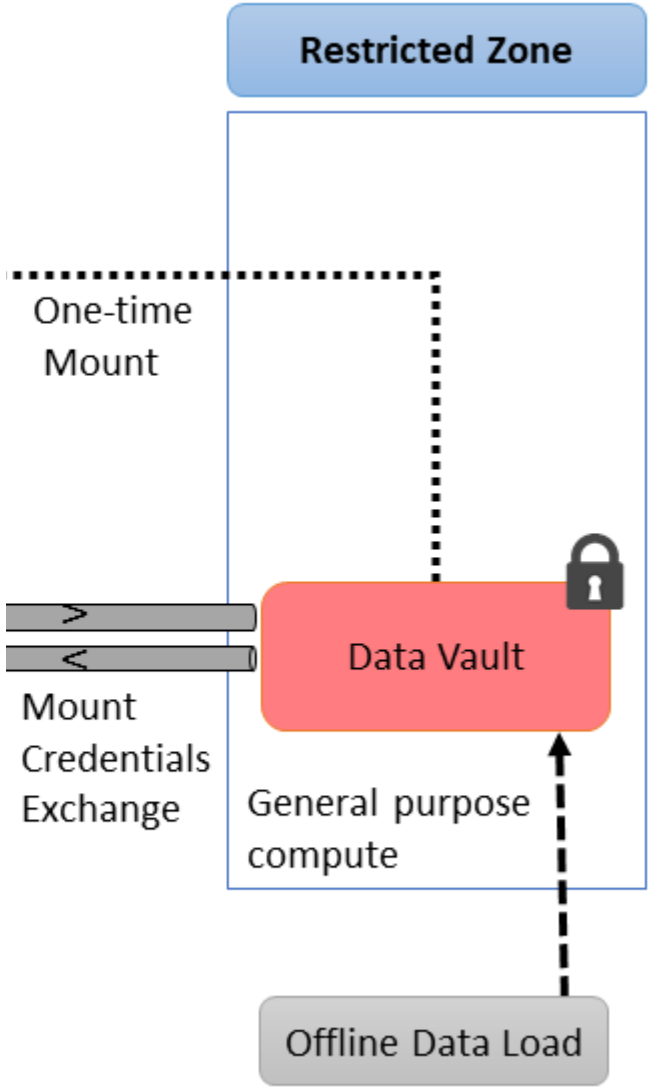

Public zone
Consumer Node
Code
Results
web, ssh
ssh
Signed Code
Vetting Node
Run results,
Audit
General purpose
compute
Offline Signing
One-way Queue
Manual, offline process
Remote Access



One-way Queue >

Manual, offline process <- - ->

Remote Access ———

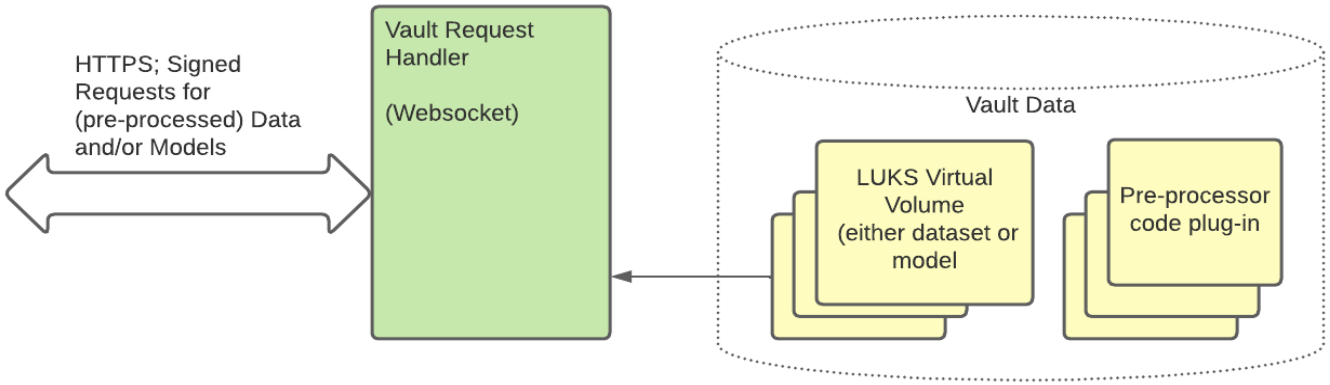


One-way Queue
Manual, offline process
Remote Access

Outcomes & Lessons Learned

Work continuing on second iteration

- Need to integrate domain expertise in workflow
- Protect models as well as data
- Federation – plurality of collaborators
- Integration with operational and research workflows
- Hardware Security for key management etc.
- Research into automated vetting
- Paralleism; websocket data interface



Conclusion

Towards Airlock 2.0

- Controlled access to large ‘eyes off’ datasets
- Future is interoperable; scalable; federated
- Data and model ‘recipes’
- Interest from a variety of problem domains

eResearch Australasia
2023 Conference
16 - 20 October
Brisbane Convention & Exhibition Centre
AeRO

AiLECS lab
Ai for law enforcement and community safety

MONASH University