

# Preliminary Implementation Recommendations: A National Approach to Genomic Information Management

## A National Approach to Genomic Information Management

This is a time of unprecedented momentum in global health genomics, with growing demand for clinical genomic testing globally and rapid acceleration of investment in genomic research<sup>1</sup>. Internationally, Australia is celebrated as an early adopter of health genomic innovation, however the application of genomic technologies in research and clinical settings brings unique challenges in technical and data infrastructure. With these rapidly expanding genomic datasets, **Australia is imminently facing a crisis in genomic information storage, analysis and sharing**, and is at least five years behind other genomic national initiatives in coordinated approaches to health data management<sup>2</sup>.

The Australian <u>Blueprint for a National Approach to Genomic Information Management</u> (NAGIM) is a digital genomics blueprint developed for the Project Reference Group on Health Genomics, under the National Health Genomics Policy Framework Implementation Plan<sup>3,4</sup>. The Blueprint delivered principles and guidelines from evaluating the current jurisdictional, operational, and technical landscape in Australia. The NAGIM Blueprint concluded that a **standards-based approach, using a federated cloud or hybrid model**, would be the most appropriate strategy for a national approach to genomic information management in Australia. The Blueprint also included a proposed roadmap of activities, as relevant to data infrastructure, genomic medicine, governance, and genomic research.

## NAGIM Prototype Program

Australian Genomics was tasked by the Australian Government Department of Health to develop recommendations for implementing the Blueprint and progressing the national approach to genomic information management for Australia. To inform these recommendations, prototypes for genomic data infrastructure were launched in 2021 to develop and encourage adoption of scalable, interoperable and extensible approaches to the collection, storage and use of genomic data in Australia.

Genomic data infrastructure stakeholders nationally were invited to participate in prototype development, in an open call leveraging existing capabilities/funding. They were tasked with addressing the priority areas identified in the Blueprint, with the goal of identifying the best combination of components that can serve as the basis for long-term national research infrastructure.

In alignment with the NAGIM Blueprint recommendations for infrastructure, prototypes were required to address:

- 1. Federated frameworks
- 2. Standards-based approaches
- 3. Interoperability across systems
- 4. Alignment with international data sharing initiatives
- 5. Cloud-based or hybrid solutions

 $<sup>{}^1\,</sup>https://www.australiangenomics.org.au/publications/australian-genomics-business-continuity-plan/$ 

 $<sup>^{\</sup>rm 2}$  Australian Genomics Domestic and International Data Infrastructure Surveys, 2020

<sup>&</sup>lt;sup>3</sup> https://queenslandgenomics.org/national-approach-to-genomics-information-management/

 $<sup>\</sup>label{eq:product} {}^4 \ https://www1.health.gov.au/internet/main/publishing.nsf/Content/national-health-genomics-policy-framework-2018-2021$ 

Nine prototypes were developed over five months, building components of the proposed NAGIM research ecosystem (Figure 1). Prototypes were completed and presented in December 2021.





An international expert panel was convened to advise and evaluate the NAGIM Blueprint Implementation prototypes in January 2022<sup>5</sup>.

The three 'critical pillars' required for NAGIM-enabled infrastructure are:

**Interoperability:** able to interact with other systems to seamlessly exchange and make use of information;

**Scalability**: able to accommodate growth and manage / process large-scale datasets; and **Extensibility**: able to accommodate new additions and expand functionality.

### **NAGIM Implementation Recommendations**

This report outlines the preliminary implementation recommendations for a National Approach to Genomic Information Management for Australia, informed by the international panel's NAGIM prototype evaluations, and parallel stakeholder consultation.

A comprehensive report and roadmap will be released mid-2022.

<sup>&</sup>lt;sup>5</sup> https://www.australiangenomics.org.au/projects/progress-the-implementation-of-the-nagim-blueprint/

# **Summary of Preliminary Implementation Recommendations**

#### Recommendation 1: Governance, Coordination and Workstreams to Progress the NAGIM Roadmap

Appoint a coordinating body, governance and working group(s) to manage, monitor and report on the parallel work streams established to progress a national approach to genomic information management in Australia.

#### Recommendation 2: Architectural Foundations for an Integrated Australian NAGIM Infrastructure

Commence building the architectural foundations, and the core genomic data management infrastructure, for a federated NAGIM research ecosystem. This should include progressing foundational 'End-to-End' (ETE) genomic platforms and integrations between them, delivered as a pilot to inform design of the production-level operations and capabilities.

#### Recommendation 3: Clinical Information Systems, Priorities, and Intersections

Establish alignment and intersections of the NAGIM research ecosystem with Australian jurisdictional clinical systems, and progress clinical priorities. This should include national regulatory frameworks for clinical data sharing, minimum requirements for clinical data collection, standards for security and privacy for progressing federated clinical settings, and clinical infrastructure pilots.

#### Recommendation 4: Data Governance and Ethical, Legal and Social Considerations

Identify and address the governance and legal requirements, and ethical and social considerations, associated with a federated NAGIM ecosystem. This should include national frameworks and agreements for data sharing; management of culturally sensitive collections and Indigenous data sovereignty; and critically, consideration of data donor consent (research, clinical), autonomy and privacy.

#### Recommendation 5: Key National Services for NAGIM

Establish the key national services required to operate a federated NAGIM ecosystem in Australia. This should include an Australian digital researcher identity and access service, and national user support services.

#### Recommendation 6: Data Standards and Harmonisation

Achieve sector agreement to establish the data standards, data harmonisation processes and data ingestion procedures to be adopted across the federated NAGIM ecosystem. This should include use of standards developed through the Global Alliance for Genomics and Health (GA4GH), HL7, and other international standards.

#### **Recommendation 7: Production Operations**

Establish the strategy and requirements for successful transition of pilot NAGIM infrastructures to production platform(s), and an operational federated NAGIM ecosystem. This should include competitive tender for mature system-level operations; consolidation of academic and industry partner networks, and a long-term sustainability model.

#### Recommendation 8: Innovation in NAGIM

Invest in a sustained strategy of innovation with evaluation and incorporation of new technologies to apply to the NAGIM ecosystem. This should be progressed separately to core data management and user support services, with a dedicated budget to support future proofing of the ecosystem to new technology and data demands.

Each of the above Recommendations should be progressed with independent but interrelated workstreams as tabulated below:

## Table 1 – Proposed Workstreams and Deliverables to Support the NAGIM Implementation Recommendations

WORKSTREAM	DELIVERABLES	DEMONSTRATED NAGIM EXEMPLARS*	KEY STAKEHOLDERS
NAGIM Governance and Coordination (WS1)	<ul> <li>Establish governance framework and implementation strategies;</li> <li>Expand stakeholder networks;</li> <li>Report on progress against agreed activities and milestones.</li> </ul>	Australian Genomics	State/Territory and Commonwealth Government and portfolio agencies; clinical and research health data infrastructure providers and users
Architectural Foundations (WS2)	<ul> <li>Assess and map the existing Australian genomic data infrastructure landscape;</li> <li>Plan the establishment of a national research data commons platform;</li> <li>Build integrations between the data commons and cloud-based ETE platforms;</li> <li>Pilot a federated approach to national data sharing using international standards;</li> <li>Develop a national strategy for a multi-cloud Australian ecosystem;</li> <li>Assess the role of HPC and on-premises data storage;</li> <li>Develop strategies and systems to bridge individual institutions into the ecosystem.</li> </ul>	Prioritised ETE platforms from the prototyping evaluation <sup>6</sup> by the: Australian BioCommons / UMCCR; and Garvan Institute of Medical Research	The NAGIM community of practice; Australian Research Data Commons (ARDC); commercial cloud providers; AAF
Clinical Information Systems and Priorities (WS3)	<ul> <li>Engage with clinical, diagnostic and health system stakeholders;</li> <li>Progress a national regulatory framework and agreements to support health data sharing;</li> <li>Establish national standards / minimum requirements for clinical and genomic data capture (phenotype, pedigree, metadata, consent, genomic data, patient identifiers);</li> <li>Build capability for diagnostic labs to access data nationally for primary clinical activities (variant data; genotype-phenotype data for interpretation; aggregated genomic data or control data for clinical pipelines);</li> <li>Establish security and privacy standards to safeguard data, in pursuing interoperable infrastructure and federated data access.</li> <li>Undertake a cross-jurisdictional pilot of clinical information sharing.</li> </ul>	TBD **	Australian Digital Health Agency (ADHA); Australian Institute of Digital Health (AIDH); accredited public pathology providers; State/Territory genetic services, health infrastructure platforms and digital health units.

<sup>&</sup>lt;sup>6</sup> NAGIM Prototype Assessment COLLATED 28.02.2022

Data Governance	• Establish a <b>governance body</b> for ethical, legal and social oversight of the	TBD **	Data custodians' legal and
and	NAGIM ecosystem including representative community involvement:		governance officers: policy and
Ethical, Legal and	<ul> <li>Establish an Indigenous genomic data governance strategy and advisory</li> </ul>		bioethics collaborators (GHFM 2021
Social	body for a NAGIM ecosystem:		Opportunity ELSI of Data
Considerations	Establish national governance policies processes and agreements to		Governance Consortium and
(WS4)	support a national data commons and federated NAGIM infrastructure:		Indigenous Genomics consortium <sup>7</sup> );
	<ul> <li>Assess the implications of clinical and research consent practices in each</li> </ul>		Experts in Indigenous data
	iurisdiction for NAGIM:		sovereignty and communities
	<ul> <li>Pilot solutions to enable appropriate progression of NAGIM informed by</li> </ul>		(ACCHO); NHMRC/MRFF;
	community engagement:		State/Territory and Australian
	<ul> <li>Implement hest-practice approaches to capturing representing and</li> </ul>		Government Departments of
	respecting data dopor's consent for secondary use including but not		Health.
	limited to dynamic consent		
Key National	Ruild on foundations established by the AAE to deploy a <b>federated</b>	Prioritised national	National Collaborative Research
Services	<ul> <li>Build of Houndations established by the AAF to deploy a rederated</li> <li>researcher identity and access management service.</li> </ul>	services from the	Infrastructure Strategy (NCRIS)
(\\/\\\S5)	Invitation to tonder for a national organisation to establish and operate	prototyping evaluation <sup>8</sup>	National Health and Medical
(**33)	• Invitation to tender for a national organisation to establish and operate	hy the Australian	Research Council (NHMRC)/
	Dovelop systems to support incentivise and mandate use of a NAGIM	Access Federation	Medical Research Future Fund
	• Develop systems to support, incentivise and manuate use of a MAGINI	$(\Delta \Delta F)$ .	(MREE): Universities and Medical
	• Establish the long term national data retention strategy	Australian BioCommons	Research Institutes: AARnet
Data Standarda	Establish the long term hational data retention strategy.	Driaritized standards	
Data Stanuarus	Obtain consensus on the common standards;	based tools from the	Clobal Alliance for Conomics and
Harmonisation	• Agree on a minimum data model for genomic data platforms;	prototyping ovaluation <sup>9</sup>	Hoalth: NACIM platform and
	• Support technical adoption of standards;	by the	sorvico providors
(\$\$50)	• Define standards to operate on <b>interfaces</b> between components;	Australian	service providers
	<ul> <li>Establish structured procedures for ingesting data into the NAGIM</li> </ul>	BioCommons/LIMCCP	
	platforms;	Garvan Institute of	
	• Support availability of <b>standardised clinical and phenotype data</b> in the	Medical Research	
	ecosystem.	CSIRO eHealth Centro	
		CSIRO eHealth Centre	

<sup>&</sup>lt;sup>7</sup> GHFM 2021 Grant Opportunity

<sup>&</sup>lt;sup>8</sup> NAGIM Prototype Assessment COLLATED 28.02.2022

<sup>&</sup>lt;sup>9</sup> NAGIM Prototype Assessment COLLATED 28.02.2022

Production	• Invitation to tender to identify organisation(s) to support the production-	TBD **	Research and clinical data
Operations	standard, national operations of the national data commons and		generators and users; data donors
(WS7)	associated components of the federated NAGIM research ecosystem		and communities; policy makers,
	• Establish a partner network for the development and operations of a		legislators and infrastructure
	NAGIM ecosystem;		funders; international partners
	• Commission an independent security audit and develop a national		
	security strategy;		
	• Establish the long-term sustainability model / business continuity plan for		
	the national data commons and federated NAGIM ecosystem.		
Innovation	• Establish separate funding and activity streams to support NAGIM	TBD **	Public and private research groups;
(WS8)	innovation, distinct from the foundational data management and		industry partners; CSIRO
	infrastructure;		
	• Invite expressions of interest to contribute innovative tools and methods		
	to incorporate into the NAGIM ecosystem;		
	• Progress tools and innovations that align with advances and standards		
	from international initiatives.		

\* Demonstrated exemplars from the NAGIM Prototype Program

\*\* To be determined by the future NAGIM co-ordinating body and workstreams (WS1)

# **Conclusion and Next Steps**

To progress the implementation of a National Approach to Genomic Information Management in Australia, **a coordinating entity** should be appointed (WS1) to advance the foundational data infrastructure (WS2) with a **funded NAGIM Implementation Pilot**.

**The co-ordinating entity** should be advised by the Commonwealth, State and Territory Departments of Health, associated portfolio agencies (such as the Australian Digital Health Agency), clinical and research health data infrastructure providers, and should be led by Australian Genomics, and the future national agency for genomic medicine - Genomics Australia.

**The foundational infrastructure** to be progressed in the Implementation Pilot phase should be based upon prioritised End-To-End (ETE) platforms from the NAGIM Prototype program evaluation: the Australian BioCommons/UMCCR and the Garvan Institute for Medical Research prototyped platforms.

These pilots should interface with **parallel workstream activities** progressing national services around digital identity and access management (WS5), and integrated standardised clinical phenotype systems (WS6). Coordination of and engagement with key partners to address the clinical intersections (WS3), as well as data governance frameworks and ethical, legal and social considerations (WS4), should also be initiated in the Pilot phase.

The outcomes of a NAGIM Implementation Pilot, will inform the design of the **Pre-Production strategy** and tenders to support the establishment, maintenance and sustainability of **Production Operations** (WS7) (see Figure 2 below).



## Figure 2 – Strategy for progressing NAGIM implementation

In parallel with the Pilot, the NAGIM implementation program should commence assembling the primary working groups associated with pilot workstreams [see Figure 3 below], and finalise stakeholders and resourcing requirements.





The first step to the formation of these working groups, initiation of the deliverables, and establishment of the implementation pilot is the **appointment of the co-ordinating and governance body (WS1)**. This will require an independent organisation with demonstrated success in national coordination, national stakeholder networks and interdisciplinary implementation projects for genomics; Australian Genomics has well-demonstrated capabilities in these areas and is well positioned to undertake this role.

A full report of the NAGIM Implementation Recommendations and proposed strategy for progressing NAGIM will be delivered to Government mid-year.

# Acronyms

AAF	Australian Access Federation
ADHA	Australian Digital Health Agency
ACCHO	Aboriginal Community Controlled Health Organisation
AIDH	Australian Institute of Digital Health
ARDC	Australian Research Data Commons
ELSI	Ethical, Legal and Social Implications
ETE	End-To-End [genomic data platforms]
GA4GH	Global Alliance for Genomics and Health
GHFM	Genomics Health Future Mission
НРС	High Performance Computing
MRFF	Medical Research Future Fund
NAGIM	National Approach to Genomic Information Management
NHMRC	National Health and Medical Research Council
UMCCR	University of Melbourne Centre for Cancer Research
WS	Workstream

9