CyberPath can get the framework right and still get the system wrong

CyberPath defines occupations, skills and experience well. But once its standards have economic value, people will optimise around them. Without deliberate incentive design, CyberPath risks creating signals that neither reliably validate skills and experience nor encourage people to build them.

Prepared by Benjamin Mossé for ACS, AISA and Home Affairs | 04 September 2026

Imagine CyberPath creates a technically perfect definition of a competent cyber practitioner

Every participant in the ecosystem still responds to their **own incentives**:

Agent	Rational local objective	Potential conflict with system objective
University	Enrolments, completion, reputation, funding	Pressure to make credentials attainable and scalable
TAFE/RTO	Completions, funding, employability	Pressure toward standardised, teachable evidence
Student	Employment for minimum cost/time	Optimise for the credential employers recognise
Employee	Promotion/mobility	Acquire evidence that passes HR gates
Consulting firm	Utilisation and margin	Leverage junior labour while signalling capability to clients

See Appendix B for more examples

Nobody needs to behave badly
for the system to produce a bad outcome

Each agent can behave perfectly rationally and ethically according to their local objective—and the emergent outcome can still undermine CyberPath's national objective.

Problem Statement

According to the National Strategy, employers face a mismatch between job requirements and employee skills, insufficient job-ready experience, and lack assurance that qualifications are fit for purpose.

Desired State

Clear standards to validate cyber skills and experience.

Missed Insight

CyberPath is not merely a workforce framework. Once adopted, it becomes an economic signalling system.

See Appendices A-C

Source: Australian National Cyber Security Strategy 2023-2030

CyberPath's biggest threat isn't bad standards. It's Goodhart's Law.

What will rational actors do once CyberPath has economic value?

→ employers begin requiring it
→ candidates optimise for it
→ educators teach toward it
→ assessment providers standardise it
→ employers increasingly trust the credential
→ economic value attaches to passing
→ pressure grows for cheaper/faster preparation and assessment
→ proxies emerge for the underlying capability
→ the credential and capability begin to separate

Disclaimer:
I'm not claiming we already know precisely how every actor will behave. I'm arguing that CyberPath cannot responsibly assume they won't respond to the incentives it creates.

A security control isn't secure because its specification is good.
It's secure if it remains effective when intelligent actors have incentives to defeat it.

Before CyberPath progresses to pilot, require every major mechanism to pass a behavioural and incentive stress-test

For every important CyberPath actor:

1. Identify what they optimise for
2. What CyberPath changes economically for them
3. How they can legitimately game the signal
4. What emergent behaviour results
5. What mechanism would realign their incentives

Without this stress-test, CyberPath risks recreating the exact problem it was designed to solve: employers still unable to trust that recognised skills and experience reflect genuine capability.

The goal is not just a better framework.
It is a system in which genuine capability is more rewarding than the appearance of capability.

What happens to Australia's cyber workforce strategy if CyberPath becomes widely adopted but weakly correlated with actual capability?

Appendix A – The ontology that was underdeveloped in CyberPath's design

The current CyberPath ontology is sophisticated about the *objects* in the workforce system (i.e., occupations, roles, domains etc.).

The Capability Framework explicitly says it is trying to recognise genuine professional competence rather than relying simply on qualifications, theoretical knowledge or job titles.

But where are the agents, their incentives, and the relationships between them?

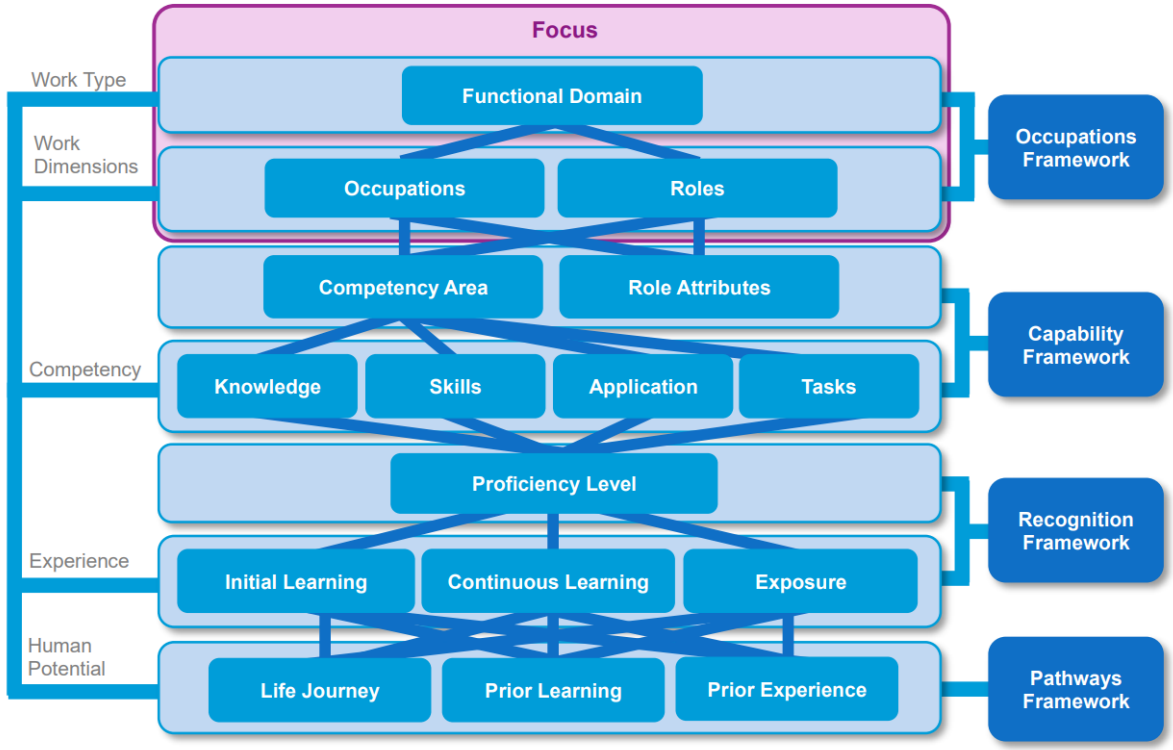

Figure 1 CyberPath Ontology

Source: CyberPath Occupations Framework

Appendix B – How participant's own objectives might conflict with the National Strategy's stated objectives

Agent	Rational local objective	Potential conflict with system objective
Employer	Fill vacancies quickly and cheaply	Use credentials as inexpensive screening proxies
Certification provider	Credential adoption/volume	Lower friction increases addressable market
HR/recruitment	Process candidates efficiently	Favour legible proxies over expensive capability assessment
Regulator/ government	Adoption, participation, policy outcomes	Favour accessibility and uptake
Experienced practitioner	Recognition of genuine capability	Wants difficult-to-fake experience to retain value

Table above is a continuation of slide 2

Appendix C - What specific CyberPath design choices become different once you account for agent behaviour?

The test for every design choice:

If this has economic value, how will rational actors optimise around it—and will that optimisation increase genuine capability or just improve the appearance of capability?

Agent-based analysis will change the framework architecture.

	Without agent analysis	With agent analysis
Design assumption	"Workplace evidence is stronger than qualification evidence."	"Employers optimise for lowest cost, highest confidence signals."
Risk if assumed	Employers may not require workplace evidence.	Employers revert to cheap credential proxies.
Design choice	Accept self-declared workplace experience with light validation.	Make high-integrity evidence economic and easy to consume.
Likely behaviour	Employers use credentials as screening proxies.	Employers use and trust verified high-integrity evidence.
Consequence	Credential and capability separate over time.	Signals remain aligned with genuine capability.
Stronger design choice	Maintain framework hierarchy.	Design for incentive-aligned use and verification.

Illustrative example

Appendix D – Three levels of evidence for testing how CyberPath will behave in the real world

Map agent incentives and likely system behaviours

Simulate how agents may respond to CyberPath

Scientifically test whether CyberPath produces the intended outcomes

Conceptual

Modelled

Empirical

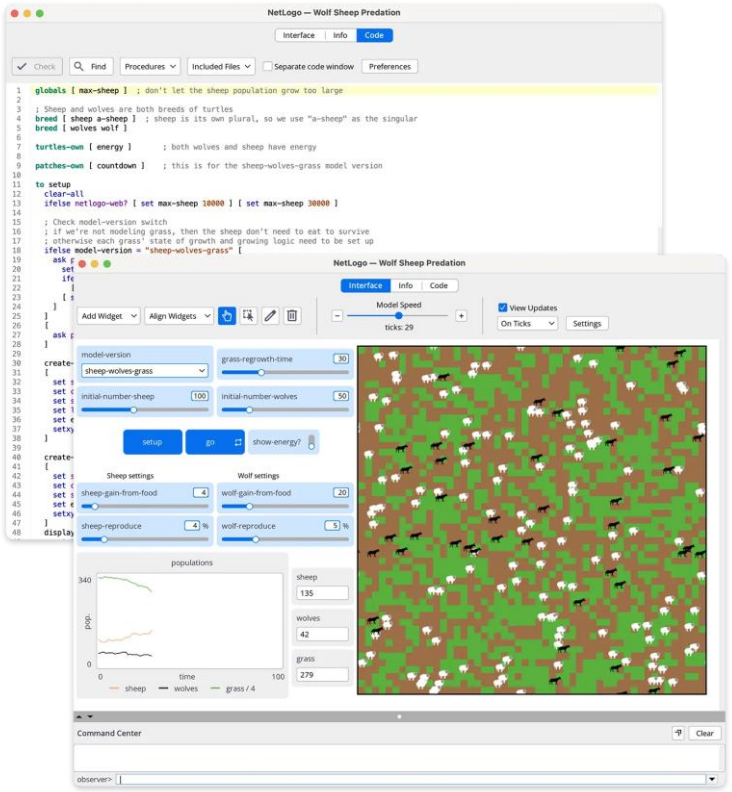

Image: NetLogo ABM software

A $1.9M public investment deserves evidence that CyberPath will work as a system—not just as a framework.