



ARGUS ACUMEN + NATIONAL-SCALE COMPUTE

Independent acceptance testing — and legacy-code porting and modernization — for scientific software on national-scale HPC

THE VALIDATION PROBLEM

Two realities collide in scientific HPC. AI agents now produce sophisticated software fast — but generation and acceptance are different functions, and the agent's own say-so isn't independent verification. Meanwhile the labs still run on decades-old scientific codes — much of it aging Fortran — with C and other decades-old scientific languages alongside it — hard to staff, and risky to change. And every new machine — each exascale or GPU architecture — forces a port, with the same question every time: did it change the answer? All of it too consequential to retire, rewrite, or re-port on faith — and it needs the same thing: evidence-backed proof the code still does what it must.

WHAT ARGUS ACUMEN DOES

Argus Acumen — powered by OneDroid Argus — brings independent verification and validation (V&V) to scientific software — proving a code does what it must, with evidence. **For new work**, the coding agent builds without access to the acceptance scenarios Argus will use — a holdout test; Argus then tests the finished build against those unseen scenarios and returns a verdict backed by cryptographic evidence that can be replayed and independently reviewed — separating builder from verifier. **For existing code**, Argus turns the same machinery inward: it analyzes legacy codes to understand them, surfaces and rationalizes long-standing bugs, and — when a code is modernized into a current language — proves the new implementation matches the original on feature completeness, numerical accuracy, and performance.



PROPOSED NATIONAL-SCALE VALIDATION

PWW proposes a controlled Argus evaluation on a national-scale supercomputing environment, using real lab codes chosen by the labs' own scientists — not a synthetic demonstration. Two complementary tracks:

- **Acceptance under HPC conditions** — can software meet independently defined acceptance criteria under the computational and operational conditions of scientific HPC? Argus runs the holdout scenarios and produces a replayable evidence record.
- **Legacy analysis & modernization** — take a set of lab-chosen hero codes; use Argus Acumen to understand them, find and rationalize their bugs, then port or modernize one or more — to a new architecture (GPU/ARM) or a modern language — and let Acumen prove the result against the original for feature completeness, numerical accuracy, and performance. Verification & validation (V&V) you can stand behind.

Each engagement fixes the workload, acceptance criteria, test boundaries, execution environment, and evidence requirements before validation begins.

WHAT THE EVALUATION IS DESIGNED TO ESTABLISH

- **Independent verification** — separate software generation from acceptance.
- **Representative HPC validation** — exercise real scientific codes in a national-scale compute environment.
- **Legacy de-risking & modernization** — analyze aging codes, surface latent bugs, and prove modernized rewrites equivalent to the originals.
- **Architecture-port confidence** — prove a code moved to a new GPU/exascale machine still produces the original's answers.
- **Predefined acceptance criteria** — define success before execution, not from the builder's own assessment.
- **Replayable cryptographic evidence** — preserve verdict evidence for independent review and replay.
- **Repeatable methodology** — a validation pattern that extends to more codes, workloads, and environments.

WHY PWW + ARGUS

Providentia Worldwide combines AI architecture, high-performance computing, and mission-critical software engineering — the rare mix needed to both validate AI-built systems and rescue the legacy scientific codes the labs have relied on for decades. OneDroid Argus makes that work independent, evidence-backed, and repeatable.

PROPOSED NEXT STEP

With a host lab: select representative national-laboratory codes and a target compute environment; agree on acceptance criteria, execution boundaries, and the evidence package; then run a focused Argus validation — acceptance, legacy analysis, or both — and review the results with stakeholders.