A proof-governed autonomy system for career conversion in embodied creative research and development

Alan N. Pham[1,*]

[1]AO Labs; institutional affiliation and coauthor list to be confirmed before submission

[*]Correspondence: Alan Pham

Autonomous agents are increasingly able to plan, monitor, and execute tool-mediated workflows, but a career objective is normally treated as motivation rather than as a measurable control problem. We report an early proof-governed system that converts a target role, Walt Disney Imagineering Research and Development mechanical Imagineering in Glendale, into a logged state machine with role-fit dimensions, guardrails, daily cycles, evidence counters, and a weekly manuscript update. The system is intentionally bounded: it does not fabricate credentials, send applications, claim relationships, or perform sensitive outreach without human approval. At the first evidence lock, the system tracked six role-fit dimensions, four portfolio anchors, one daily cycle, zero proof logs, zero outreach events, and one published weekly paper snapshot. The current fit score was 61 of 100, with the dominant bottleneck being principal-level network signal. The contribution is therefore not a claim that the system has secured a job. It is a falsifiable protocol for making an otherwise ambiguous career-conversion process auditable, adaptive, and constrained by truthful evidence.

1 Introduction

Tool-using language-model systems can now reason, call external services, write artifacts, and update deployed software (*1–4*). These capabilities create an opportunity to treat long-horizon personal objectives as monitored systems rather than as episodic intentions. A career target is especially difficult because the desired outcome depends on evidence quality, technical fit, social trust, timing, applications, and human judgment. Without a state model, repeated effort can feel intense while remaining scientifically uninterpretable.

This paper describes an early deployment of an autonomous career-evidence system aimed at WDI Research and Development mechanical Imagineering roles. The system uses the active WDI Research and Development Imagineer - Mechanical Design Engineer role as the live rung and the Principal R&D Imagineer - Mechanical Engineer profile as the north-star target (*5*). The positioning line is deliberately stable: mechanical PhD, soft robotics, creative prototyping, and AI-assisted tools for human-facing physical experiences.

The design borrows from proof-governed revenue experimentation: define a measurable state, choose the current bottleneck, constrain allowed actions, record interventions, and update a paper from real evidence rather than from aspiration. The system is not allowed to invent credentials, inflate relationships, send spam, or apply on behalf of the operator. It can monitor state, update dashboards, maintain a paper, identify the next constraint, and create or improve public artifacts that are grounded in existing work.

2 System architecture

The first deployment consists of a public dashboard, a Railway-hosted backend, a JSON runtime state file, and scheduled automation. The backend exposes health, operations-check, journal, weekly-paper, event, and daily-cycle endpoints. The dashboard is intentionally minimal: it shows backend state, fit score, bottleneck, weekly paper status, evidence counters, active experiment, next action, role-fit dimensions, and recent log entries.

3 Decision policy

Table 1. State variables at the first evidence lock. Values are reported from the live system after the first daily-cycle and weekly-paper update.

Variable	Recorded value
Target company and location	Walt Disney Imagineering R&D, Glendale, California
Live rung	WDI Research and Development Imagineer - Mechanical Design Engineer
North star	Principal R&D Imagineer - Mechanical Engineer
Fit score	61 of 100 after one daily cycle
Dominant bottleneck	Principal-level network signal
Evidence counters	Zero proof logs; one daily cycle; zero outreach events; four portfolio anchors
Active experiment	WDI proof packet v0
Human approval boundary	Applications, sensitive outreach, referrals, and claims involving real people

The policy scores six role-fit dimensions: mechanical depth, creative prototyping, human-facing physical experience, principal-level network, Glendale application packet, and the paper/system itself. The next action is chosen from the weakest current signal. At the first lock, the weakest signal was principal-level network, so the selected intervention was to create one real review path. This means preparing an artifact for a reviewer and asking a specific fit question, not sending automated messages or pretending a relationship exists.

This policy is intentionally conservative. It can make public artifacts, update the dashboard, refresh the manuscript, and record evidence. It cannot manufacture the social proof that a principal-level role requires. If an action requires a real person's attention or reputation, the system must stop at preparation and request approval before sending.

4 Results

The current deployment is functional but early. The dashboard and backend are live, the weekly paper endpoint can publish a snapshot, and the site exposes a PDF manuscript. The first daily cycle increased the fit score from 60 to 61 by recording the selected action and improving the leadership-network dimension from 34 to 39. This is useful as an instrumentation result, not as a job-market result.

5 Limitations and guardrails

Table 2. Role-fit dimensions after the first autonomous cycle. Scores are internal operating metrics, not external hiring probabilities.

Dimension	Score	Current interpretation
Mechanical depth	80	Strong existing base, but still benefits from visible calculations and design artifacts.
Creative prototyping	82	Strongest signal; should remain attached to concrete physical demonstrations.
Human-facing physical experience	72	Credible but needs clearer guest-facing translation.
Principal-level network	39	Current bottleneck; requires real review paths and trust signals.
Glendale packet	50	Needs a compact, role-specific application surface.
Autonomous career system	44	Early but now instrumented through dashboard, endpoints, and weekly manuscript.

The system cannot guarantee an offer. A job decision depends on timing, hiring needs, competition, interview performance, portfolio interpretation, and institutional context. The useful claim is narrower: the system can make progress legible, choose the next bottleneck, prevent fake progress, and turn repeated effort into evidence. It should be evaluated by proof velocity, reviewer-ready artifacts, quality of feedback paths, application readiness, interviews, and eventual conversion.

The guardrails are part of the result. The system forbids fabricated credentials, fake outreach, invented Disney relationships, automated applications, and claims that are not supported by public or logged evidence. These constraints reduce short-term aggressiveness, but they preserve interpretability. If the system ever reports strong progress, that progress should be traceable to artifacts, logs, reviewers, applications, or outcomes.

6 Availability

The live dashboard is available at `https://imagineer.aolabs.io/` once DNS resolves, with a fallback at `https://aolabs.io/imagineer/`. The live backend is available at `https://imagineer-app-production.up.railway.app/api/imagineer/ops-check`. The weekly paper endpoint is available at `https://imagineer-app-production.up.railway.app/api/imagineer/weekly-paper`. This manuscript is an initial public artifact and should be refreshed from logged evidence as the system accumulates real proof.

References

1. S. Yao *et al.*, ReAct: Synergizing reasoning and acting in language models. *International Conference on Learning Representations* (2023).

2. T. Schick *et al.*, Toolformer: Language models can teach themselves to use tools. *arXiv preprint arXiv:2302.04761* (2023).

3. Q. Wu *et al.*, AutoGen: Enabling next-gen LLM applications via multi-agent conversation framework. *arXiv preprint arXiv:2308.08155* (2023).

4. L. Wang *et al.*, A survey on large language model based autonomous agents. *arXiv preprint arXiv:2308.11432* (2023).

5. The Walt Disney Company, *WDI Research and Development Imagineer - Mechanical Design Engineer* (2026; `https://www.disneycareers.com/en/job/glendale/wdi-research-and-development-imagineer-mechanical-design-engineer/391/93733641696`).