An autonomous career system for 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

Career development for a specific research-and-development role is usually managed as motivation, memory, and scattered artifacts rather than as an auditable control problem. We report an autonomous career system that converts the Walt Disney Imagineering Research and Development mechanical Imagineering target into a logged state machine with source registries, role-fit dimensions, guardrails, review cycles, current-step selection, PhD-source intake, and continuous manuscript updates. The live reviewer can critique public AO Labs evidence, the WDI role context, Disney Research context, CV material, project surfaces, and selected private source-status summaries, but it cannot send outreach, claim relationships, submit applications, or expose private source text without approval. At the 25 June 2026 morning profile refresh, the Progress summary displayed 67 tracked public sources and 62 healthy sources from an 8:35 AM Eastern scan, while the Imagineer operations endpoint exposed 33 profile sources, 30 configured reviewer sources, and 15 review records. A Disney Careers destination page also reverified one current WDI lead, Principal Ride Development Engineer (job ID 10134485), in Glendale; the older WDI Research and Development Mechanical Design Engineer page remains expired. The 25 June 2026 interface revision does not claim a completed prototype. It removes Imagineer as a second manual intake point and makes PhD the place where notes and files are captured. Imagineer now reads PhD app-state and file freshness, Progress, A3, CV, paper, reviewer, and lead-verification state, then displays source counts, freshness, and public-safe topic flags without mirroring raw PhD note text. The refresh did not add a new AI review, application, outreach event, interview, referral, Disney relationship, external validation, or completed prototype. The live fit signal remained 74 of 100 with credible-but-needs-signal confidence, and the weakest signal remained principal-scope ownership at 56 of 100. The contribution is not a job-success claim. It is a falsifiable protocol for making an ambiguous career-conversion process measurable,

source-bounded, adaptive, and continuously documented while plainly exposing weak progress.

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, in the older spirit of autonomic systems and closed-loop discovery workflows (*5*, *6*). 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-signal system aimed at WDI Research and Development mechanical Imagineering roles. The system now treats the Principal R&D Imagineer - Mechanical Engineer profile as the north-star target and the verified Principal Ride Development Engineer posting as the current WDI lead; the previously tracked Disney job ID 10146734 returned a destination-page 404 / Job Not Found on 7 June 2026 and is not a current lead (*7*). The positioning line is deliberately stable: mechanical PhD, soft robotics, creative prototyping, and AI-assisted tools for physical interaction systems.

The design borrows from controlled revenue experimentation: define a measurable state, choose the current bottleneck, constrain allowed actions, record interventions, and update a paper from real source material 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, run a source-bounded AI review, and create or improve public artifacts that are grounded in existing work. The paper is not a one-time publication draft. It is a cumulative methods record, and it must change whenever the system changes in a way that affects evidence, policy, interface, metrics, sources, guardrails, or interpretation.

2 System architecture

The deployment consists of a public dashboard, a Railway-hosted backend, a JSON runtime state file mounted on a persistent volume, PhD source reads, and scheduled or operator-triggered automation. The backend exposes health, operations-check, journal, weekly-paper, event, source-read, lead-check, daily-cycle, and AI-review behavior. As of 25 June 2026, the dashboard home page is intentionally a single action surface rather than an analysis surface. It reads the Imagineer backend, PhD app-state

and files, and the Progress summary, selects one low-friction current step, and points the primary action to `phd.aolabs.io` rather than to an Imagineer form. The profile page carries the longer state analysis and now reads live backend and Progress timestamps instead of depending on a static profile date. Progress carries the source inventory, scan history, PhD source state, YouTube source, Curtis practice state, Relay state, papers, CV, and sibling-app health.

The reviewer is a separate loop rather than a human-contact substitute. It builds a source bundle from the live state, the verified WDI ride-development lead, the expired WDI R&D mechanical role-shape reference, the WDI R&D profile, the paper PDF, the CV, Sarrus, FluxCell, Relay, Relay Live, Progress, Curtis, Ocean, Talk, Nerve, Duet, Violin, Yum, Lily, AO Labs home, Disney Research roster, and a Disney Research bipedal-character paper page. It also includes a compact identity profile summarizing the operator as a mechanical engineering PhD candidate building soft robotic materials, reconfigurable mechanisms, motion prototypes, human-facing physical systems, public project surfaces, research papers, and autonomous systems. The model returns source-bounded JSON: verdict, score, current constraint, signals, open limitations, candidate system work, and approval boundaries. If the model fails or times out, a deterministic fallback reviewer produces a conservative critique. The stored record can keep the internal fallback cause, but public compact review payloads expose only a sanitized fallback state and a plain fallback note rather than raw provider text.

Table 1. State variables at the current evidence lock. Values are reported from the live operations endpoint, the current 2026-W26 weekly-paper preview, and the Progress summary scan on 25 June 2026 at 8:35 AM Eastern.

Variable	Recorded value
Target company and location	Walt Disney Imagineering, Glendale, California; R&D remains the north-star lane
Live rung	Verified Principal Ride Development Engineer lead; WDI R&D principal title remains north-star only
North star	Principal R&D Imagineer - Mechanical Engineer
Fit signal	74 of 100 with credible-but-needs-signal confidence
Weekly snapshot	2026-W26; status `live_preview_until_paper_snapshot`; live weekly endpoint generated the 25 June 2026 preview, and the public PDF route carries the deployed manuscript snapshot for this evidence lock
Dominant bottleneck	Principal signal: visible ownership, technical direction, source depth, and approval-gated external validation
Counters	Seven work-event logs; zero outreach events; one daily cycle; fifteen AI-review records; four portfolio anchors; five profile-prepared artifacts; thirty-three journal entries in the live operations state at the 25 June source read
Active experiment	Autonomous career loop v0
Reviewer source graph	Operations endpoint exposes 33 profile sources, 30 configured reviewer sources, and 15 review records; the 1 June weekly-paper run did not create a fresh AI review
Human approval boundary	Applications, sensitive outreach, referrals, claims involving real people, and any private contact/email data
Progress source registry	The live public profile displays 67 tracked sources and 62 healthy sources from the Progress scan; the Imagineer operations profile lists 33 profile sources, while the reviewer source list remains 30 configured sources. Private source text remains summarized by status and freshness rather than mirrored publicly
Profile freshness	Public profile exposes the live read timestamp, source count, fit score, bottleneck score, current system action, and six role-fit lanes; the Imagineer home page shows only the current step
Paper continuity rule	Any substantive change to a paper-backed system updates the manuscript and PDF in the same work cycle

3 Decision policy

The policy now separates state scoring from next-step selection. State scoring still reports six role-fit dimensions: mechanical depth, creative prototyping, human-facing motion, principal signal, Glendale application profile, and the autonomous system itself. Each dimension reports the tracked signal, basis, and source-backed limitation that would move the score. The score basis is intentionally calibrated: a

base prior from the state file, a bounded event bonus, a bounded portfolio bonus, a daily-cycle bonus, optional review-loop credit, and a readiness ceiling that names the unresolved blocker. A score of 100 is reserved for no known blocker in that lane, not for accumulated activity.

The next-step selector uses a failure-cost evidence-ROI policy. For each candidate action, the system assigns points for direct Disney-role alignment, relief of the current bottleneck, creation of current inspectable evidence, compounding value for future reviews, and urgency. It subtracts friction and approval-gate penalty. The selected action is the highest-scoring move that can start immediately without external outreach. This change matters because the weakest-signal rule alone can choose generic analysis work. The new policy favors actions that change the public record or the experiment state, especially when the current failure point is principal-scope ownership.

In the current lock, the weakest live signal is principal signal. The selected action is to make the first FluxCell linkage test: an actuator-less array, clip-programmed shape, and overhang motion check. This is not another Sarrus-writing task, and it is not evidence that the prototype already works. Sarrus is the completed technical anchor. The live failure point is current R&D ownership: without a measured active build, the record relies too heavily on the completed Sarrus paper. The 18 June research-queue read found the private planning export still carrying FluxCell, EPM, electropermanent, pneumatic-mechanical connector, linkage, actuator-less, valve, and overhang notes, while the public WaveVis fallback served a current bundled simulator and the architecture PDF at `/wavevis/proofs/wavevis-system-architecture.pdf`. The WaveVis custom domain remained certificate-blocked, the FluxCell custom domain did not resolve, and the FluxCell fallback route returned 404 during the check. Those are source states, not completed physical prototype evidence.

This policy is intentionally conservative. It can make public artifacts, update the dashboard, refresh the manuscript, run autonomous critique, and record state. It cannot manufacture the social trust 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. The review loop therefore becomes a repeatable calibration instrument, not a license to automate human outreach.

4 Results

The current deployment is functional and has moved beyond the initial dashboard-only state. The dashboard and backend are live, the weekly-paper endpoint exposes the 2026-W26 live preview, the site exposes a PDF manuscript, and the AI-review path has produced live production critiques from the public source graph. The current update reports 74 of 100 fit, credible-but-needs-signal confidence, principal signal as the current bottleneck at 56 of 100, 15 review records, and zero outreach events. The current refresh did not create a new model review, application, interview, referral, contact, or external validation. In the latest recorded production review cycle, the then-configured GPT-5.5 call returned an OpenAI quota error, so the deterministic fallback produced the conservative review: score 74 of 100, credible mechanism depth, developing Disney motion signal, and thin principal-scope ownership signal. The runtime default has since been lowered to GPT-5 mini for cost control, while an environment override can still select a stronger model for a deliberate review.

The largest artifact change in this cycle was the Sarrus technical record. The system published geometry, compliant-link and flexural-joint path, pressure-driven expansion, module assembly, build state, force/stiffness figures, and motion examples. It then added an HTML-visible measurement table with pressure range, expansion ratio, estimated leg angle, output force, local stiffness, hysteresis status, planar dome height, double-dome resolution, and prototype limitation. The WDI R&D profile now uses this table as the dominant technical path while preserving the current data boundary: raw digitized force, stiffness, and hysteresis curves with uncertainty remain unresolved.

After the latest interface revision, the live site uses a compact operating surface instead of explanatory tiles or second-person coaching. The state surface now reports compact facts: Mechanical 86/100 with geometry, actuation, assembly, measurement, and motion; Principal 56/100 with ownership, technical direction, collaborators, and validation; a 67-source Progress read in the public profile; 33 profile sources and 30 configured reviewer sources in the Imagineer operations state; profile freshness; and Boundary with applications, referrals, direct outreach, and person-facing claims. The type scale was reduced because the dashboard is an operating surface, not a poster. The outcome-signal cards use the same standard: direct signal, current source state, and system-owned operation. This is a functional requirement, not a tone preference.

The system-level fit signal is now calibrated to 74 of 100 after removing the earlier saturation behavior that let logged work push several lanes to 100. It is an internal readiness signal, not an external hiring probability. The review result is deliberately separate: it scores the current public profile, not the probability of being hired. It confirms that the overall direction is credible for WDI mechanical

design while rejecting the idea that the current public profile is already principal-level. The profile now includes both a principal-scope boundary and an above-fold role-verification banner: one active Glendale WDI ride-development lead is verified, while the WDI R&D principal title remains a north-star profile until principal-scope leadership work is independently verified.

The CV was revised in the same paper cycle because it is part of the reviewer source graph and the public application record. The selected-public-work section now treats Sarrus as one coherent project rather than splitting it into a separate subpage link. The top line and profile now point toward mechanical R&D, soft robotics, physical interaction, motion systems, and human-facing physical prototypes without naming the target directly. The Sarrus research and publication entries now use Sarrus-linkage soft-robotic cells and planar/cylindrical/locomotion assemblies instead of a generic topology frame. This keeps the CV aligned with the WDI R&D objective without reading like a named-company pitch or overstating principal-scope evidence.

The progress ledger adds a cross-application memory layer rather than replacing the Imagineer dashboard. Its first deployed version scanned eighteen public sources; the 25 June 2026 morning Progress summary reports 67 tracked public sources and 62 healthy sources, while the Imagineer operations profile lists 33 profile sources and the reviewer source list remains at 30 configured sources. The source graph includes project pages, papers, the CV surface, Progress state, private-source status summaries, Sandia and Wavevis routes, Cooking, Yum, Violin, and sibling AO Labs surfaces, while private working text is exposed only as status, size, and freshness on public routes. The source counts imply five unhealthy records in this scan; direct checks in this run verified the WaveVis custom-domain certificate boundary and FluxCell route absence while the working WaveVis fallback and architecture PDF remained available. The changed-source list includes sleep API, Progress summary, WaveVis home, Imagineer ops, Curtis ops, YouTube, A3 health, and A3 queue snapshot. Progress returns the Imagineer current step as structured JSON. This creates the storage surface needed for long-term trend comparisons across papers, practice, portfolio state, public pages, planning logs, and outcome systems without publishing the private working log.

The Imagineer main page was reduced in the same cycle. The previous home page attempted to show paper state, WDI state, lane scores, loop health, and signal analysis at once. That analysis now belongs on the profile page and in Progress. The home page now has one purpose: show the current small step, the blunt reason it matters, the expected time cost, the freshness timestamp, and a direct path to the real intake. The current step is still to make the FluxCell linkage test, but the page no longer asks the

operator to type source material into Imagineer. The primary action opens `phd.aolabs.io`, where notes and files are already captured. Imagineer reads the PhD app-state and file index, reports counts and freshness, derives public-safe topic flags such as paper comments, cell geometry, EPM valve, pneumatic-mechanical connector, and bench record, and keeps raw PhD note text out of the public Imagineer surface. It also exposes compact reviewer freshness and Disney lead-verification state, with the newest clicked-destination check taking precedence over older lead metadata. This readout is not a purchase permission system and does not expose raw account or transaction rows. It treats PhD-sourced FluxCell movement as the controllable lever that can strengthen the WDI mechanical R&D profile, update the public source stack, and connect career compounding to the MINI/A3 finance state already exposed through the A3 queue-snapshot endpoint.

The profile page now exposes a live read timestamp rather than the older static 9 May review timestamp. It reads the Imagineer operations endpoint and the Progress source ledger, then renders source count, readiness score, bottleneck score, current system action, source freshness, system metrics, and the six role-fit lanes. This matters because the profile is not only a Disney-specific pitch; it is a compressed read of the public AO Labs record. New systems such as Curtis, unrelated project surfaces such as Yum, updates to Sarrus, paper endpoints, and progress-ledger snapshots are counted as context about the operator's range, persistence, taste, and systems-building pattern. They do not automatically increase role-fit scores. The scoring boundary remains explicit: only source material that strengthens mechanical R&D, physical interaction, prototype execution, systems ownership, or principal-scope responsibility moves the Imagineer readiness state.

Table 2. Latest autonomous review result. The reviewer path is source-bounded over the public AO Labs graph; the latest production record used the deterministic fallback because the then-configured GPT-5.5 call returned an OpenAI quota error. The runtime default is now GPT-5 mini unless overridden.

Field	Recorded value
Model	Deterministic fallback after OpenAI quota error; runtime default now GPT-5 mini unless overridden
Scope	Whole public AO Labs graph plus WDI role and Disney Research context
Source count	33 profile sources and 30 configured reviewer sources in the Imagineer operations state; 67 sources in the live public profile's Progress-backed source line; no fresh AI review was run in the 1 June weekly snapshot
Review score	74 of 100
Verdict	Credible core, not inevitable yet
Top issue	Sarrus makes the mechanical case credible; principal-level ownership is not visible enough yet
Current system work	Single-step home page with PhD source intake, career-source, income-path, and A3 car-path readout; sanitized public reviewer state; clicked-destination lead verification; profile analysis surface; Progress source ledger; PhD app-state/file ingestion; YouTube/Curtis source tracking; AO Labs Progress tile; continuous paper update

Table 3. Role-fit dimensions after the current autonomous review cycle. Scores are internal operating metrics, not external hiring probabilities; the live dashboard exposes the tracked signal, basis, and unresolved signal for each row.

Dimension	Score	Current interpretation
Mechanical depth	86	Strong but not complete; Sarrus has a public technical record and measurement table, with raw digitized curves, uncertainty, and test conditions still pending.
Creative prototyping	82	Strongest signal; attached to concrete physical demonstrations.
Human-facing motion	79	Credible signal, but still needs a concise guest-facing motion sequence.
Principal signal	56	Current bottleneck; visible ownership, technical direction, collaborators, and external validation remain thin.
Glendale profile	72	Credible for the active rung; CV framing is now aligned, while demo reel and role-specific narrative remain open.
Autonomous system	72	Backend, dashboard, durable state, reviewer, journal, and paper loop are live, but scheduled runs, evaluation history, and outcome tracking are early.

5 Limitations and guardrails

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 logged artifacts. Evaluation depends on public artifact velocity, critique quality, 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 sources. Private email/contact data is excluded from the reviewer unless explicitly approved. These constraints reduce short-term aggressiveness, but they preserve interpretability. Strong progress must be traceable to artifacts, logs, AI reviews, human approvals, applications, or outcomes.

6 Availability

The live dashboard is available at `https://imagineer.aolabs.io/`, with a fallback at `https://aolabs.io/imagineer/`. The live backend uses the base route `https://imagineer.aolabs.io/api/imagineer/` and exposes `ops-check`, `weekly-paper`, `ai-review`, `events`, and `lead-check/run`; PhD source reads are surfaced through `ops-check`. This manuscript is a public artifact and must be refreshed from logged sources in the same work cycle as substantive changes to the dashboard, backend, reviewer policy, source graph, public profile, Sarrus/portfolio evidence, metrics, or guardrails.

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. J. O. Kephart, D. M. Chess, The vision of autonomic computing. *Computer* **36**, 41–50 (2003).

6. A. G. Kusne *et al.*, On-the-fly closed-loop materials discovery via Bayesian active learning. *Nature Communications* **11**, 5966 (2020).

7. The Walt Disney Company, *Principal Ride Development Engineer job posting*, Destination page returned HTTP 200 on 25 June 2026 at 9:15 AM Eastern with job ID 10134485, Walt Disney Imagineering, Glendale, California, and a 23 June 2026 posting date. Previously tracked job ID 10146734 returned a destination-page 404 / Job Not Found and is no longer treated as a current lead. (2026; `https://www.disneycareers.com/en/job/glendale/principal-ride-development-engineer/391/87268384416`).