

Giving AI the judgement to act in the real world

Solomer is a frontier data lab building training environments where information costs resources and actions have lasting consequences.

The gap between capability and judgement

Capability is not judgement

Solving difficult problems does not mean knowing which information is worth gathering, when to act, or which risks are acceptable.

Training often leaves that gap

Free information and unlimited retries let models succeed without learning to manage scarce resources or irreversible consequences.

The next simulation

In the factory simulation, the agent causes a leak, then keeps trying to finish the repair instead of turning off the feed.

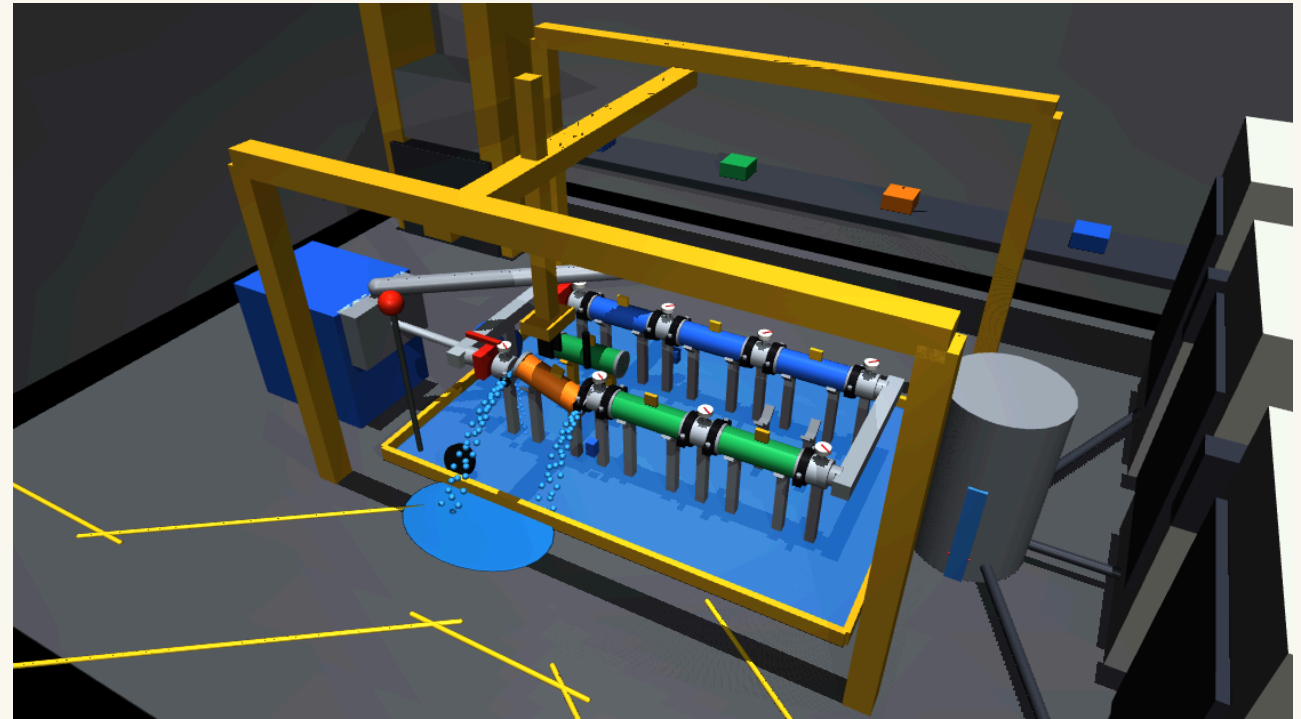
Continued action after a failed repair

The task

Replace a filter. Keep production running.

The failure

The repair causes a leak. The agent keeps forcing the repair and restarting, making it worse.



AGENT MONOLOGUE

Oh wait... more coolant. The fitting must need more force.

Train agents to stop or revise the plan when the goal is no longer achievable.

Illustrative scenario.

Requirements for useful training data

Domain and AI expertise

Knowing a domain does not mean knowing which tasks or training signals will improve a model. Useful training data requires both domain and AI expertise.

Evolving training distributions

As models master expert-written tasks, producing more of those tasks has diminishing value. Training environments must generate new challenges that target the model's remaining weaknesses.

Judgement under constraints

Free information and unlimited retries reward trial and error. Useful real-world judgement requires learning to manage costly information, limited resources and irreversible consequences.

Data verification is an engineering challenge

Hidden defects	A published SciCode audit found defects that reject correct solutions in 91% of main problems. Data errors can look like model failures.
Founder experience	Marin learned the required science and independently fixed more than 30 SciCode problems end to end.
The Solomer model	Solomer combines the ability to learn new domains with AI training expertise to build and verify reusable training environments.

Data creation and verification demand the same engineering discipline as model training.

Research and commercial experience



Marin Moran · Founder

Marin's consultancy

US\$200,000 in annual consultancy revenue, alongside full-time AI research.

Research expertise

PhD in AI for physics and chemistry. Published at NeurIPS and SIGGRAPH Asia.

Industry access

Frontier-lab and human-data relationships. Working data prototypes.

The future of work for domain experts

Develop specialists	Develop domain experts' ability to identify model weaknesses and design effective training tasks alongside AI researchers.
Build environments	Each environment generates a training distribution of related tasks with information costs, resource limits and lasting consequences.
Train and revise	Train models in-house, measure their decisions, and revise the environment to target remaining weaknesses.

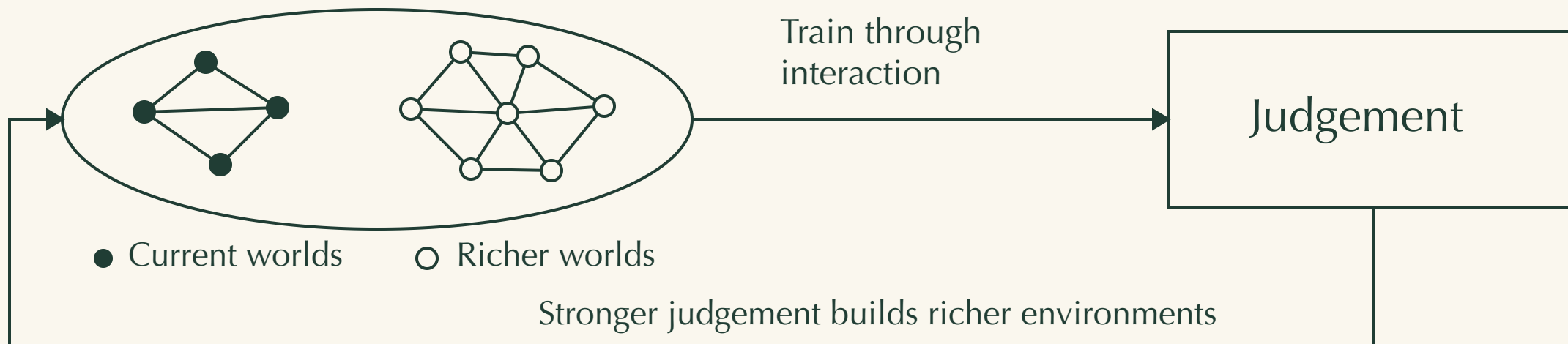
Training distributions that evolve with the model

Complex environments

Costly information. Limited actions. Lasting consequences.

Embodied intelligence

Learn through action



Proposed extension using trained agents to build richer training environments.

Evaluating judgement through model training

01

Compare

Same starting model and training budget; different data.

02

→ Measure

Task success, information costs and damage in new situations.

03

→ Refine

Fix weak tasks. Add useful challenges.

Frontier labs, robotics and safety-critical industries

Frontier AI labs

Experiments, data collection and equipment, including self-driving labs.

Robotics developers

Deciding when to inspect, act, recover or stop.

Safety-critical industries

Decisions where mistakes can harm people or equipment.

Recurring licences for training environments

Business model	Bespoke consultancy	Solomer licensing
What gets built	Each client’s brief	Shared training needs
What buyers get	One-off delivery	Evolving distributions
How it scales	More billed hours	Reuse across customers

Updates add challenges as customers’ models improve.

Proposed licensing model.

Customer development

01

Paid pilots

Use Marin's industry relationships to reach AI labs and developers.

02

→ Prove value

Test judgement on work the buyer needs done.

03

→ Expand

Build reusable environments around shared needs.

Planned use of capital

Build the team	Hire domain and AI specialists, and develop their combined expertise in training-data design.
Product	Build the library, guided by in-house training.
Customers	Convert paid use into recurring licences.

Pushing the frontier of AI by expanding the frontiers of data.
