



OSTERI LABS - MISSION CARD 01

Build a Tool That Travels

Facilitator's guide

A group exercise about building a research institution under uncertainty.

30-45 MINUTES / GROUPS OF 3-5 PEOPLE / ONE SHARED SCREEN

Part of “Trade-offs, Events, and Trajectories: Research Institution Simulator” series

What this is

This is intended as a facilitation device rather than a prediction engine. The value is in the discussions that it will spark in the team.

Participants found a research institution with one mission: build a tool that works reliably, meets a real user need, and can be used beyond the founding team.

The group sets founding traits, chooses a governance model, and allocates starting resources for their institute. Those choices shape options and vulnerabilities, but they do not determine a single outcome. Random events will occur throughout the run that force the group to reallocate resources, make personnel decisions, and refocus the institute towards specific projects. Probabilities for these events are based on their founding selections.

Session at a glance

GROUP SIZE

3-5 people per shared run. Split larger groups into teams of 3-5 and compare reports afterward.

TIME

30-40 minutes: about 10 minutes to found the institution, 15 minutes to run it, and 15-25 minutes to debrief.

MATERIALS

One internet-connected screen per team. Ask one person to capture the group's reasoning.

BEST FOR

Groups discussing research strategy, institutional design, or program leadership.

Before people arrive

- Open labs.osteri.blog and test the simulation to familiarize yourself with how it works.
- Based on expected participants, choose one shared institution or several parallel teams.
- Decide how choices will be made: consensus, vote, or a designated director.
- Plan for sufficient time for the final debrief.

RUNNING THE SESSION

A suggested run of show

- **Frame the mission - 3 min.** Read the mission aloud. Explain that the group is designing an institution to tackle this specific problem, not selecting their ideal general institution.
- **Found the institution - 8-12 min.** Allow the group to discuss the six traits, governance, and allocation. Ask them to consider what each choice protects and what it makes harder.
- **Run the simulation - 15-30 min.** The cycle will play in real time with critical decision points along the way. The groups can use Skip to Next Decision when they understand the system or want to leave more time for discussion.
- **Read the report - 5 min.** Have the group read the mission, institution, legacy outcomes, founding design, and key decisions.
- **Debrief - 15-25 min.** Have the group explain the rationale for their founding design and key decisions as well as the outcome. This is also a good point to test counterfactuals and compare the simulated institution with real ones.

If time is short

Use Skip to next decision after the first cycle and reserve at least 10 minutes for debrief. A rushed run with a real discussion is better than a complete run with none.

FACILITATOR STANCE

Do not explain every causal relationship in advance. Give participants enough information to make reasoned bets. Part of the fun is making decisions under uncertainty.

MECHANICS

What's happening under the hood.

The following pages (up to pg. 10) review the different mechanics at work in the simulator and how they interact with each other.

The founding design

Six traits define recurring pressures throughout a run. In the game engine, none of these traits act like an isolated bonus or have a universally correct setting. Multiple strategies can lead to success, but their likelihoods are different.

Founding traits

RESOURCE PATIENCE

How long funders tolerate weak visible progress. Patience protects long-horizon work; impatience makes weak individual cycles more consequential.

FUNDING CONCENTRATION

Whether support comes from many funders or one dominant sponsor. Diversification reduces dependence but adds coordination; concentration increases exposure to a funding shock.

RESEARCHER DISCRETION

How much room teams have to redirect work. Discretion helps pursue anomalies and user needs, but can create conflict with the mandate or founder dependence in the wrong governance setting.

PERSONNEL CONTINUITY

How long expertise remains inside the institution. Continuity builds memory and reduces turnover risk, while increasing recurring commitments and making restructuring harder.

PROBLEM-USER COUPLING

How closely researchers work with users. Closer coupling improves fit and opportunity capture, while consuming slack and coordination capacity.

PORTFOLIO BREADTH

How many technical routes remain in play. Breadth preserves alternatives; focus preserves attention, runway, and slack.

MECHANICS

Governance and Founding Allocation

Governance selects ordinary project work automatically. These are calculated by the game engine with the team's initial design influencing the probabilities. The team will intervene when a consequential situation exceeds those ordinary rules.

Governance

DIRECTOR-LED

The founder selects projects quickly. Coordination is easier, but succession and founder dependence can become severe.

MANAGED AUTONOMY

Teams propose work and leadership prioritizes mission-relevant cases. It balances direction with local judgment.

DISTRIBUTED SELECTION

Projects advance through collective review. Defensible evidence is favored; broad portfolios create more selection friction, while departures are easier to absorb.

RESEARCHER AUTONOMY

Researchers choose among feasible projects. Exploration increases, along with coordination costs and the chance of drifting from an impatient sponsor's mandate.

Founding allocation

BALANCED

No pronounced specialization across runway, capability, memory, and user access.

PEOPLE-HEAVY

More execution capacity and memory, with higher recurring commitments.

EQUIPMENT-HEAVY

Stronger platform and validation work, with less financial flexibility.

RESERVE-HEAVY

Longer runway and more slack, with limited early technical capacity.

USER-NETWORK-HEAVY

Stronger user access and field learning, with less engineering depth.

MECHANICS

What happens in a cycle

Each cycle turns the institutional design into a set of project choices, results, costs, events, and possible interventions.

- **Select current work.** Governance ranks eligible projects and activates up to three. Breadth, allocation, user coupling, discretion, and unmet evidence change those rankings.
- **Resolve project results.** Each active project produces a setback, informative negative, expected result, strong result, or breakthrough. See more in “Project outcomes” below.
- **Update the institution.** Evidence, capability, memory, slack, runway, team health, and stakeholder confidence respond to the work.
- **Draw events.** A personnel event and an external event may occur in the same cycle. Slack can absorb setbacks or capture an opportunity.
- **Pause or continue.** A consequential decision, mission completion, financial failure, or the end of the horizon may stop ordinary play.

Project outcomes

SETBACK

No evidence gain. Available slack contains some damage; without it, team health suffers.

INFORMATIVE NEGATIVE

No immediate evidence gain. The failed route becomes institutional memory and improves later work on that project.

EXPECTED RESULT

Incremental evidence and a small capability gain.

STRONG RESULT

A larger evidence gain that also strengthens scientific confidence.

BREAKTHROUGH

A rare, unusually large advance. It helps, but cannot bypass evidence dependencies.

DECISIONS UNDER UNCERTAINTY

The interface shows qualitative conditions and likely trade-offs. Exact probabilities remain hidden, so decisions stay bets rather than calculations.

MECHANICS

How project results are calculated

The engine separates project selection from project outcomes. Founding choices can change what the institution works on, how likely that work is to succeed, or how costly shocks become.

Each active project receives a quality score based on the institution's capabilities, memory, slack, governance, founding allocation, and previous results. The relevant factors differ by project: user relationships matter more for pilots, while technical capability matters more for prototyping and validation.

Higher quality makes strong results and breakthroughs more likely. Lower quality makes setbacks and informative negative results more likely.

Randomness is a part of the fun, so a well-designed institution can still struggle, and a weak one can occasionally succeed.

Examples from the engine

- Permanent personnel arrangements reduce annual turnover risk from 19% to 9%.
- Learning from an informative negative slightly improves that project's future odds.
- Funding concentration increases the damage caused by a funding shock, but not how often shocks occur.
- Strong results advance evidence faster but cannot bypass mission prerequisites.

READ THE NUMBERS CONDITIONALLY

Because several factors interact before the final probability is calculated, the same change may have a larger or smaller effect in different institutions.

MECHANICS

Mission completion

The mission is complete only when the tool works, is validated, meets a real need, and can travel beyond the founding team. There are four mission evidence objectives in the interface that show progress.

FUNCTIONAL EVIDENCE

Unproven -> principle demonstrated -> repeatable prototype -> robust in relevant use.

VALIDATION

Untested -> internal checks complete -> independently validated -> reliable across settings.

USER EVIDENCE

Untried -> user need established -> workflow fit demonstrated -> demonstrated user value.

TRANSFER PREPARATION

Tacit and team-bound -> working documentation -> external reproduction -> transfer-ready.

These objectives interact with and are dependent on each other.

- User discovery can begin immediately; deeper pilots require a working tool.
- Functional evidence can reach a repeatable prototype but cannot become robust in relevant use without meaningful user evidence.
- Validation depth is limited by functional maturity.
- Documentation can begin early, but transfer cannot be completed until functional, validated, and useful evidence are mature.
- A breakthrough accelerates eligible work. It does not erase prerequisites.

Current work

The Current work panel shows the projects governance selected for that cycle. Results from those projects may advance the linked evidence track. It is a portfolio rather than a second progress score.

Available work includes principle and measurement, prototype platform, calibration and reproducibility, user discovery and pilots, documentation and transfer, and a later exploratory branch for high-discretion institutions.

MECHANICS

The institution has its own trajectory

Mission evidence and institutional health can diverge. For example, a tool may improve while the organization becomes financially or politically fragile.

Institutional capacity

RUNWAY

How long current reserve can cover the institution's commitments.

CAPABILITY

The ability to execute planned technical work. It shapes core technical and validation outcomes.

MEMORY

Knowledge that survives individuals. It strengthens validation and transfer and softens personnel loss.

SLACK

Capacity beyond planned work. It absorbs setbacks and makes unexpected opportunities actionable.

TEAM HEALTH

The institution's morale and internal stability.

Stakeholder confidence

FUNDERS

Shapes continuing support and the severity of renewal pressure.

RESEARCHERS

Shapes project performance and the institution's ability to retain energy behind the work.

USERS

Shapes user-facing project outcomes and reflects whether useful evidence is accumulating.

WATCH FOR DIVERGENCE

The most interesting runs are often not clean wins or losses. Ask when mission progress, runway, and confidence begin moving in different directions.

MECHANICS

Events, decisions, and endings

Some pressures are scheduled; others become possible only when the institution's history makes them salient.

Events

- Personnel events include routine turnover and strong hires. Personnel continuity, memory, documentation, and governance shape their frequency or severity.
- External events include funding shocks and opportunities. Funding concentration changes exposure; slack determines whether an opportunity can be captured.
- A personnel event and an external event can occur in the same cycle.

Consequential decisions

The model contains five scheduled decision points, including a formative technical choice, institutional reviews, a personnel decision, and a late-stage strategic choice. Financial distress or early recognition can add conditional decisions.

The institution cannot reverse its identity every cycle. Decision descriptions state likely consequences, not guaranteed outcomes.

Founder departure

Early success may bring the founder an outside offer. If the founder leaves, the player becomes an observer and governance completes the run. Memory, continuity, governance, and accumulated founder dependence shape the disruption.

How a run ends

- **Mission complete:** the tool satisfies all four evidence requirements. The institution may continue or dissolve after successful transfer.
- **Horizon reached:** the mission remains incomplete, but the institution may still be viable and preserve useful capability.
- **Financial failure:** the institution closes or is absorbed. Its legacy depends partly on whether knowledge became transferable.

FACILITATING

Keep the conversation going

Your job is to surface reasoning, discussion, and disagreement rather than steering the group toward the configuration most likely to win.

Here are some example questions you can use at different phases of the run.

During founding

- What kind of failure is your design trying to prevent?
- Where are you accepting dependence or rigidity?
- What will this institution be unusually good at? What will it be unusually bad at?
- Do any choices appear to reinforce or contradict each other?

At a decision

- What does the group know now, and what remains uncertain?
- Are you spending any institutional resources trying to protect another?
- Which stakeholders would prefer each option: funders, researchers, users, or the founder?
- Would your stated governance in the institutional design phase actually make this choice?

If agreement comes too easily

- Ask one person to steelman the case for the rejected option.
- Ask whether the group is optimizing for mission completion, institutional survival, or transfer.
- Return to the founding profile. Is the group acting consistently with the institution it designed?

DON'T TRY TO RESCUE THE RUN

A strained, incomplete, or dissolved institution can produce a better discussion than a clean mission completion.

AFTER THE REPORT

Debrief

Use the final report as a shared object. Explain the outcome before judging it. This phase can be done in any way you choose, but here are some suggestions.

Three rounds

- **Explain.** Which founding trade-off and which decision best explain the trajectory? What was mostly luck?
- **Counterfactual.** If you changed one founding choice or one consequential decision, what outcomes would you expect to change along with it?
- **Transfer.** Where does this resemble a real institution? Where does the analogy break?

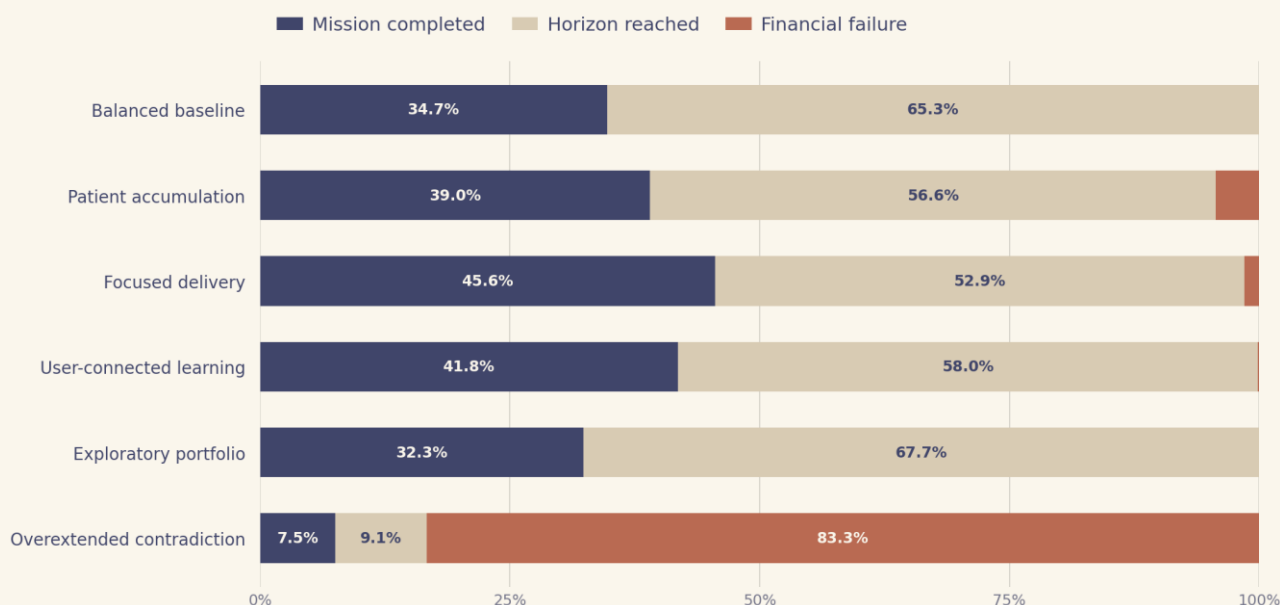
Prompts

- When did the mission and the institution's incentives begin to pull apart?
- Did the group protect visible progress, long-term capability, or organizational survival?
- Which stakeholder gained influence over the run?
- Did useful negative results become memory, or merely delay progress?
- Was founder departure a failure of succession, a test of durability, or a rational choice?
- Was the final outcome a success? Success for whom?

SEED-RUN ANALYSIS

What different strategies tend to produce

The same founding design can produce many histories. Each strategy below was run from the same set of 10,000 starting seeds and the same mission-directed decision policy, with the founder staying when offered another role.



What this suggests

- Focused delivery completed the mission most often (45.6%), but it did not dominate every seed.
- User-connected learning completed 41.8% of runs with almost no financial failures.
- Exploratory portfolios usually remained viable but more often reached the horizon without completing the mission.
- The deliberately overextended design ended in financial failure in 83.3% of runs.
- Among successful runs, every strategy had a median completion in Year 14. Design changed whether the institution crossed the line more than how quickly successful runs finished.

MODEL NOTES

What this model claims

The simulation makes institutional hypotheses discussable. It does not estimate the future of a real organization.

Limits and Boundaries

- Relationships are stylized and directional.
- Qualitative bands summarize hidden numeric states; they should not be read as measurements.
- Each run produces one possible history. A single run cannot rank institutional designs.
- Mission completion is not the only defensible outcome. Durable capability, transfer, and responsible dissolution can also matter.
- The model omits many real forces, including field-specific science, labor markets, regulation, and wider institutional ecologies.

Play the simulation at labs.osteri.blog. Read more at osteri.blog.