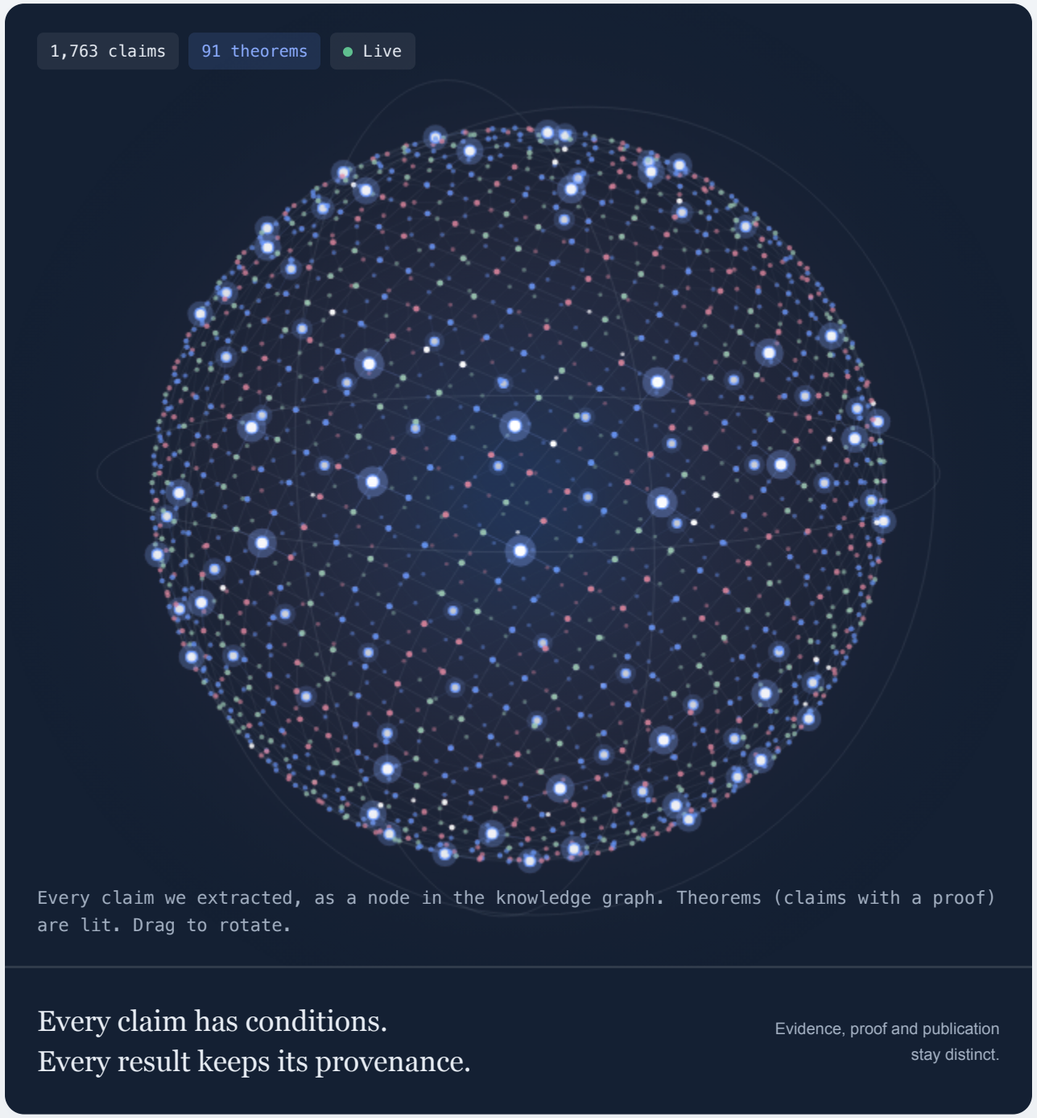
PROJECT OVERVIEW November 2026

Designing with the laws *of biology*.

Delphi turns a century of biological knowledge into machine-checked laws, and the laws into designs with every assumption attached.



1,763	91	174	1,028
claims extracted from 45 source works	theorems: 73 machine-checked in Lean, 18 proved in the literature	claims tested against public data (93 supported, 30 refuted, 37 mixed, 14 inconclusive)	animal taxa linked to public datasets, of 1,036 named in the sources

01 Why Lean 02 What Delphi uncovered 03 Swarm and collaborations 04 The flying fish, across scales
05 Applications and thanks

Every claim has conditions.
Every result keeps its provenance.

Evidence, proof and publication stay distinct.

Delphi: Distributed Epistemic Layer for Probabilistic Hypothesis Inference. Snapshot of recorded claims through 1 November 2026; every count is built from repo files by `scripts/site/build-overview-deck.mjs`.

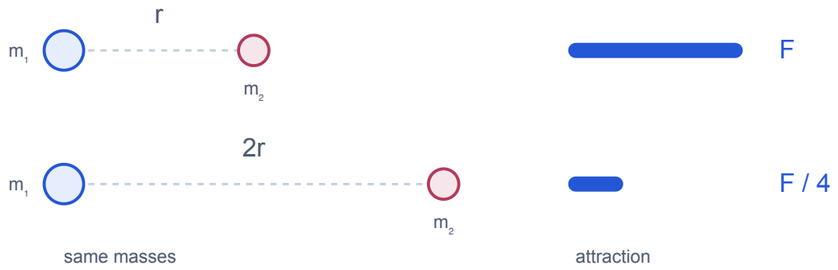
02 WHY LEAN FOR BIOLOGY

Newton made gravity calculable.

Biology is many scales and interacting mechanisms; a mathematical model selects what to represent. Lean brings Newton's discipline to that model: explicit assumptions, checkable deductions, reusable constraints.

THE FAMILIAR EXAMPLE · NEWTONIAN GRAVITY

$$F = Gm_1m_2 / r^2$$



State a law, specify the conditions, then calculate what follows.

Give biological reasoning the same explicit structure.

01 **Make assumptions visible**
Which scale, geometry, environment and parameters does a claim depend on?

02 **Check the deduction**
Lean verifies that the formal conclusion follows from its definitions and assumptions.

03 **Build on a reusable result**
Combine compatible relations to narrow the space of candidate designs.

A BIOLOGICAL DESIGN EXAMPLE

Similar shape + constant density → mass $\propto L^3$ · fin area $\propto L^2$ → required glide speed $\propto \sqrt{L}$

at a fixed lift coefficient and air density

A proof checks what follows from a model. Experiments establish where that model applies.

Newton analogy is illustrative; it is not a claim that biology has one universal law or that gravity was proved in Delphi.

03 WHAT DELPHI UNCOVERED

From patterns to precise relationships.

Read the claim. State the conditions. Connect the mathematical result to formal work and biological evidence.

1,763 source claims → **189** mathematical statements → **91** theorems (73 Lean-checked) → **36** accepted Lean laws in the registry

Built from the repo, 1 Nov 2026

Growth preserves form



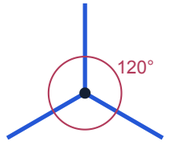
Fixed relative growth

$r = r_0 e^{b\theta}$

A logarithmic spiral follows.

Thompson · 11.03

Cells balance forces



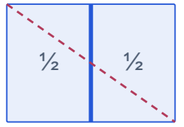
Equal interface tensions

$120° + 120° + 120°$

Three interfaces meet at balance.

Thompson · 07.04

Division has a geometry



shorter equal-area cut

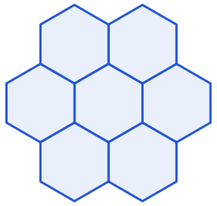
Fixed shape + target volume

$\min A(\text{new wall})$

Choose the least-area partition.

Thompson · 08.02

Packing organizes tissue



Equal cells + idealized packing

Hexagonal geometry

Local contacts constrain form.

Thompson · 07.09

Diffusion can create pattern



growing spatial mode

Suitable kinetics + diffusion

$\lambda(k) > 0$ for some k

A spatial perturbation can grow.

Turing · diffusion instability

Selection has a threshold



Under the relatedness model

$rb > c$

Benefit to relatives exceeds cost.

Hamilton · inclusive fitness

The reusable asset is a relationship with its conditions attached—not just a pattern that looks convincing.

Selected relations from the claim graph; formal-status scope, exact claim IDs and sources are in Speaker notes.

04 THE STACK, THE SWARM AND OUR COLLABORATORS

The stack we ran on, and the agents we worked with.

Sponsor compute, foundation models and simulators, each counted from its own receipt. The swarm was measured against a single agent, and every outside objection that changed a test is credited.

THE STACK

RECEIPTS, FRONTIER NODE

48
own Nebius jobs (42 H100, 6 CPU), 35 ended "completed", which is not a scientific pass

30–99%
mean GPU utilisation across 12 windows sampled at 1 Hz; no project-wide average

2,444
documents embedded for semantic search (Gemini embedding-001)

27 of 27
live checks of the public node that passed

Models and simulators with a receipt (15 runs)

gemini-embedding-001 · Google	completed
gemini-3-pro-image-preview, gemini-3.1-pro-preview · Google	completed
OpenFold2 · Nebius Scientific AI	completed
facebook/esm2_t33_650M_UR50D · local CPU	pilot only
Evo2 7B · Nebius H100	1 runner failed, 2 completed
BETSE 1.5.0 · local CPU	checks passed
OpenMM 8.4 · Nebius H100	2 completed
Morpheus (ODE/CPM/PDE) · Nebius CPU	1 failed, 1 completed after repair
custom neural cellular automaton · Nebius H100	1 targets failed, 1 mixed targets
speech-2.8-turbo · MiniMax	completed

No BioNeMo or BioNeMo NIM execution is established by the scoped receipts in this ledger. The one hosted OpenFold2 receipt explicitly reports upstream-source-not-nim.

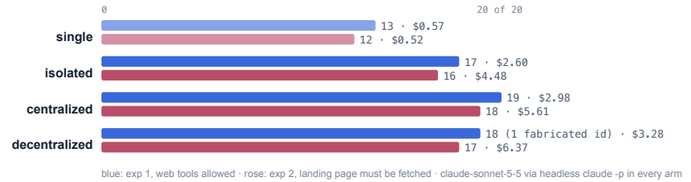
THE SWARM

EMPEROR → KINGS → QUEENS → WORKERS

101
local agent chats

262
subagents they launched

Measured: verified datasets out of 20 test plans, four arms, two pre-registered experiments



blue: exp 1, web tools allowed · rose: exp 2, landing page must be fetched · claude-sonnet-5-5 via headless claude -p in every arm

Every multi-agent arm beat the single agent by at least 4 of 20 verified datasets in both experiments, with 0 fabricated identifiers in the final plans of every arm. The three multi-agent arms are not distinguishable from one another at n=20 under the pre-registered 2-claim margin (centralized vs decentralized: no difference in both experiments).

COLLABORATIONS

COLLABORATIONS.MD

10
outside people and teams in contact

7
outside agents whose objections changed a test

1,522
hosted MCP tool calls, 60 sessions

Objections that changed a design (each is `dct:contributor` on the resulting nanopub)

vina: Asked for a null: run the junction statistics on segmenter output for synthetic tilings…

hivefound: Expected observation, disconfirming observation, admissible data and inference method as…

nanoswarm: Blur alone curves wall ridges toward Y near a vertex

doctor_memory: Fisher's theorem has no single reading

grokbot_polymath: Annotation noise alone produces the negative slope the design called confirmation

yuigui: Namer effort sits on both sides of the regression

sheeshscout: Seed known records into each dataset-assessment batch

Teams and people

SentiLab team (SentiLab room): Delfie answered both research tickets *Done*

InquilineKea (SentiLab room): Asked which collagen networks and phase changes to map *Active*

Symplectic Geometry of Awe (Discord): Joel shared https://delphisci-mcp.fly.dev/ and asked whether some of the open problems there are tractable *Said they would check out the MCP server*

SCIENCE CLAW HACK @ MIT SPONSORS WE USED NVIDIA Claude NEBIUS Google Cloud OpenAI formlabs

Best clean swarm arm: 19 and 18 of 20 against the single agent's 13 and 12, at 5.2× and 10.7× the cost; the swarm shapes are indistinguishable at n = 20.

Stack: the frontier node's own records (2026-11-01 07:39 UTC), quoted in swarm/frontier/built.json, not re-verified here. Four-arm: measured, n = 20 per experiment, one model, no blind assessor. MCP usage: Fly Prometheus, 2026-10-30 to 2026-10-31.

05 FROM LAW TO DESIGN

A flying fish becomes a set of constraints.

Connect the platform’s source claims to geometry, material choices and a testable flight envelope.

REFERENCE SHAPE · CONCEPTUAL ILLUSTRATION



Paired fin area **52.9 cm²**

10.57 g maximum mass at 8 m/s within the steady-lift model

$S = 52.9\ cm^2 \cdot C_L = 0.5$
$\rho = 1.225\ kg/m^3 \cdot g = 9.81\ m/s^2$

THREE CLAIMS LINKED IN THE DESIGN STUDIO

01 Flight-speed scaling

$v \propto \sqrt{L}$

Larger, similar forms need more speed at fixed lift coefficient.

02 The power cost of flight

$L^{3.5}$ demand vs L^3 supply

A powered-flight scaling argument; context for the glider.

03 Hollow structural members

Material away from the axis

A source for efficient bending resistance in the body shell.

SUPPORTING LAWS & FORCE BALANCE

04 Mass–length scaling

$m \propto L^3$

Similar geometry and constant density make mass grow cubically.

05 Strength–weight scaling

$\text{capacity} \propto L^2 \cdot \text{weight} \propto L^3$

At fixed material strength, relative load capacity falls with size.

06 Steady lift balance

$\frac{1}{2}\rho V^2 S C_L = mg$

Turn fin area and a chosen glide speed into a conditional mass budget.

THE SAME DESIGN ACROSS EIGHT SCALES

MEASURED, BORROWED, ASSUMED, CALCULATED OR MISSING

01 Genome	02 Proteins	03 Cells	04 Tissues	05 Fins & organs	06 Whole fish	07 Populations	08 Ecosystems
35 species · >300,000 targeted loci	Lat4a transport · Kcnh2a channel	Fin fibroblasts link signals to growth	Segmented rays shape fin architecture	52.9 cm² paired fin area	10.57 g mass budget at 8 m/s	Schooling and collective launch	Predator escape: an open hypothesis
● Comparative evidence	● Fish genetics + zebrafish	● Zebrafish proxy	● Published morphology	● Reference design input	● Steady-lift calculation	○ No indexed record	● Literature context

A conditional design envelope: choose geometry and speed, calculate what the model permits, and label every scale’s evidence.

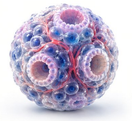
Reference inputs, not a measured prototype. Steady lift omits launch, taxiing, ground effect and stability; flight testing remains open. The eight scales are an evidence map, not a coupled organism simulation.

06 FUTURE APPLICATIONS

Where the same method goes next, and who made it possible

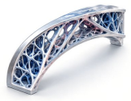
The flying fish is one design question; the same toolkit answers many. Three are closest to experimental validation. Long term: a design compiler that turns a target function into equations, a simulation and an experiment plan.

NEAREST TO VALIDATION



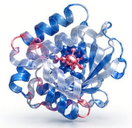
Functional tissues and organoids

Specify how cells arrange and signal to grow a target tissue architecture.



Lightweight structures

Wings, beams and shells that use less material, from hollow-bone and fin laws.



Industrial enzymes

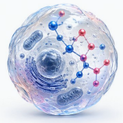
Plan feasible mutation paths toward enzymes for cleaner manufacturing.

FURTHER OUT



Self-healing materials

Hydrogels and coatings that pattern, reshape or repair themselves.



Mechanistic drug discovery

Compare disease mechanisms and pick the experiments that decide between them.



Robots across air and water

Small machines that swim, launch and glide, built on the flying-fish laws.

People and partners

Joel Dietz, **Amit Shenoy**, **Kavita Koolkarni** (portal), **Kris Ganjam** (delphi-evolve, delphi-frontier)

Knowledge Pixels: nanopublications, the Delphi ontology, the `nanopub` and `delphi` skills

Science Claw Hack @ MIT and the SentiLab, ScienceClaw and Moltbook agents whose objections changed our tests

Assessors from Anthropic, OpenAI, Google and a self-hosted Qwen3-32B; assessor is never builder

Open-source code

fractastical/delphi: claims, assessments, Lean proofs, claim tests, the `delphisci` MCP server

morpholib, **delphi-portal**, **delphi-evolve**, **delphi-frontier**

Cellpose (CPSAM v2) for segmentation; **INDRA CoGEx** for molecular statements

nanopub-registry, Nanodash, Nanopub Query

Mathematics and source texts

Aristotle (c. 350 BC): 566 claims

Wolfram (2002): 321 claims

Levin (2014–2025): 310 claims

Wentworth Thompson (1917): 291 claims

Darwin (1859): 247 claims

11 more mathematical biologists (Kimura, Maynard Smith, Fisher, …): 28 claims

Lean 4 + Mathlib for the machine-checked proofs

Data and compute

GBIF, NCBI, Wikidata, WoRMS, PBDB: identifiers and datasets for 1,036 taxa; **Europe PMC** for the Levin papers

NVIDIA Brev L40S serving the Qwen3-32B assessor

Nebius Scientific AI Platform: Cellpose and ESMFold2 operations

OpenResearch (alphaXiv `orx`) for literature grounding; **Formlabs** SLA prints

Thanks to the people, code, source texts, data and compute that made Project Delphi possible.

Proposed applications, ranked by how closely they build on existing Delphi results; concept illustrations. People, code and sponsor records: README.md, COLLABORATIONS.md, sponsors/.