



Open hope for breast cancer research.

A non-profit, open-source platform that turns validated medical knowledge into a citation-grounded AI research copilot for breast cancer.

🌐 openelpis.com · openelpis.life | ✉ hello@openelpis.com | 📍 Türkiye

Partnership & In-Kind Support Proposal — 2026

We are seeking donated cloud credits, storage, CDN, or GPU compute to run a non-profit, open-source research platform that helps breast-cancer researchers find answers faster. This document explains what we are building, why it is safe and responsible to support, and exactly what we are asking for.

1 The one-paragraph pitch

Breast cancer is the most-researched cancer in the world — and still, vital knowledge drowns in data that never reaches the people who could use it: thousands of papers a week, de-identified datasets locked in silos, and clinical insight that lives only in individual experts' heads. **OpenElpis** is a non-profit platform where verified clinicians, clinics, and labs contribute *validated* material — literature, de-identified datasets, structured findings — which then powers a citation-grounded AI research **copilot**. The model never invents; it retrieves from a curated, peer-reviewed corpus and shows its sources. Every contribution is identity-verified, expert-reviewed, and permanently provenance-tracked. It is **open source**, **non-profit**, and built so that no patient-identifiable data enters the system until governance and law fully allow it.

2 The problem we're solving

- **Knowledge is fragmented.** Answering "what's known about resistance to this targeted therapy in this subtype?" means manually crossing thousands of sources.
- **Generic AI is unsafe for this.** Off-the-shelf chatbots hallucinate citations and have no provenance — unacceptable in clinical research.
- **Good data has nowhere trustworthy to go.** Clinicians and small labs hold valuable de-identified findings with no validated, open, non-commercial home.

OpenElpis is the trustworthy middle layer: **validated in, citation-grounded out.**

3 What we're building

1. **A contribution pipeline with trust as the product.** Contributors are identity- and credential-verified (medical license, institutional email, ORCID; organizational KYC for clinics/labs). Every upload passes automated safety, de-identification, and duplicate checks, then **human expert review**, then receives an **immutable provenance record**. Nothing reaches the model without passing every gate.

2. **A validated corpus.** Only approved content: a vector knowledge base for literature (RAG), versioned datasets for structured/omics data, and a biomedical knowledge graph for reasoning about drug repurposing and mechanisms.
3. **An AI research copilot.** An open-weight orchestrator that answers researchers' questions by *retrieving from the validated corpus* and calling specialist scientific models — every claim traceable to a reviewed source. It produces research hypotheses for qualified professionals, never diagnosis or treatment advice.

4 Why this is safe — and low-risk for a sponsor

- **No raw patient data, by design, for a long time.** We start with published literature and de-identified, aggregate material only. Imaging and patient-derived data are explicitly **deferred** until de-identification pipelines, data-use agreements, and ethics/IRB oversight exist.
- **Validation, not scraping.** Human experts sign off on every accepted contribution; everything is auditable and retractable.
- **The model reads a curated corpus — it is not trained live on uploads.** No uncontrolled data flows into model weights.
- **Open by default** — open-source code, open methods, transparent data-integrity reporting.
- **Clear liability framing.** Not a medical device; gives no clinical advice — stated prominently and enforced in the product.
- **Portable across providers** — built on open standards, so your contribution is not propping up a proprietary walled garden.

5 What we're asking for

We seek **in-kind support** — donated credits, free-tier capacity, or sponsored services — to run **Phase 1** (the no-PHI foundation) and grow from there. The footprint is deliberately small and portable:

What we need	Phase 1 sizing	Why
Compute	1 instance, ~8 vCPU / 16–32 GB RAM / 200 GB NVMe, 24/7	Web app, API, review workflow, async workers, text-RAG
Object storage	250 GB – 2 TB, S3-compatible, modest egress	Validated documents, datasets, vector-store assets
CDN + WAF + DDoS	In front of the public site	A public-interest platform must stay available and secure
Managed database (optional)	Postgres, ~2 vCPU / 8 GB / 100 GB	Metadata, provenance, audit log
GPU compute (bursty)	A few hundred GPU-hours/year (A100/H100-class)	Periodic embedding generation + light fine-tuning

Equivalent retail value: roughly \$3,000–\$5,000/year for Phase 1. Any subset helps — a sponsor covering *one* of these rows materially de-risks the project.

Reading this as a specific provider?

- **Cloud / hyperscaler** — a credit grant covering compute + storage + occasional GPU.

- **CDN / edge / security** — the CDN + WAF + DDoS row, and/or zero-egress object storage.
- **GPU specialist** — the bursty GPU row, even a one-time research credit.
- **VPS / dedicated** — the 24/7 compute instance, ideally as an ongoing donated/discounted host.

6 What a sponsor receives in return

- **Visible recognition** as a founding infrastructure partner on openelpis.com, our GitHub README, and release notes.
- **A real impact story** — a published, co-authored case study on powering open breast cancer research with your platform.
- **Open-source goodwill** in the medical-AI and open-science communities.
- **Periodic impact reports** — contributors onboarded, corpus size, researcher queries served, datasets published openly.
- **Optional co-marketing** around milestones, fully at your discretion.

We are happy to sign your standard nonprofit/credits agreement and provide our legal-entity / fiscal-sponsor documentation on request.

7 Roadmap

Phase	Focus	Data risk
1 — Foundation (now)	Portal + identity verification + manual review + Tier-A literature → RAG copilot	None (no PHI)
2 — Structured data	De-identified datasets, automated de-id + schema validation, dataset versioning, reputation	Low (de-identified)
3 — Imaging & models	DICOM/WSI de-id pipeline, data-use agreements, ethics oversight, periodic fine-tuning	Governed
4 — Scale & federation	Federated learning — models travel to hospital data; raw data never leaves the institution	Minimized by design

8 About

OpenElpis is an independent, non-profit, open-source initiative founded in Türkiye. We are establishing the project's legal and governance structure (registered non-profit / fiscal sponsorship) and welcome partners — technical, clinical, and academic — who share the mission of accelerating open breast cancer research responsibly.

Contact: hello@openelpis.com · openelpis.com

This document describes a research-support platform. OpenElpis produces research hypotheses for qualified professionals and does not provide medical diagnosis or treatment advice. © 2026 OpenElpis — shared for the purpose of evaluating in-kind / sponsorship support.