# MammoChat OPTS–EGO Ledger

*Closing Breast Cancer Awareness Month with Open Data*

October 31, 2025

Dexter Hadley, MD/PhD

Founder & CEO, MammoChat™

Lake Nona Medical City, Florida | iDrDex@MammoChat.com

**Abstract**

MammoChat is a patient-centered AI platform that unites compassion, compliance, and computation to create a provenance-driven network for women's health. At its core, the **OPTS–EGO Ledger** transforms each mammogram, biopsy, and clinical note into a verifiable, patient-consented data token. Using **mCODE** (Minimal Common Oncology Data Elements), records are standardized, cryptographically notarized, and shareable under transparent consent-turning regulatory rights into real patient agency. Through a multilingual interface and community matching, the platform connects women to culturally concordant support and clinical trials. We are raising a $12 M Series A to demonstrate commercial scalability of MammoChat to validate expected improvements in anxiety, depression, and trial participation. Public deliverables, including those seeded by Florida's **$2 M Casey DeSantis Cancer Innovation Grant**, are registered as **OPTS–Grant Tokens**, linking taxpayer investment to open-source innovation and measurable patient benefit. MammoChat strives to be the first commercialized venture of a generalized framework for **AI, Blockchain and Cybersecurity in Health Care** that its founders pioneered.[1]

**Maria - 47-year-old woman**

*Maria is a 47-year-old Catholic schoolteacher from Colombia who presented after a screening mammogram revealed a unilateral asymmetry.* Biopsy confirmed **stage I luminal A breast cancer**[2–7]. She underwent **breast-conserving surgery** and adjuvant endocrine therapy without complication. After her divorce, she relocated to Florida to be closer to family but reported feeling **isolated, anxious, and fearful**[8]. Her electronic health portals were in English only, and fragmented communication added to her distress.[9–12] Social screening was consistent with **moderate anxiety** (GAD-7 = 13) and **mild depression** (PHQ-9 = 10).[13,14]

**Zaida - 52-year-old woman**

*Zaida is a 52-year-old software engineer of Pakistani heritage, an observant Muslim who lives alone and works remotely.* She presented after noticing axillary fullness, and biopsy revealed **node-positive, HER2-positive invasive ductal carcinoma**.[15–17] She began **neoadjuvant HER2-directed therapy** with plans for surgery and adjuvant radiation. Although her treatment was technologically advanced-with wearables, remote vitals, and EHR-linked data-Zaida described feeling **watched but not understood**[18]**.** Frequent alerts disrupted her work and prayer routines, amplifying a sense of **fear and isolation**.

Figure 1: Dual clinical vignettes illustrating the emotional and technological challenges faced by breast cancer survivors. Maria (left) and Zaida (right) represent contrasting experiences of isolation and anxiety-one from language and cultural barriers, the other from technology overload. Both highlight the need for empathy-native digital health solutions.[19–22]

# 1. Founder's Preface: From Awareness to Agency

**Maria and Zaida** both experienced fear, anxiety, and isolation during breast-cancer care.[23–26] Both cases reflect how breast-cancer survivors can experience profound **emotional distress** despite excellent clinical care[23,27]. Maria's challenge stemmed from **silence and disconnection** through language and cultural barriers, whereas Zaida experienced mainly **noise and intrusion** through technology overload.[28,29] For both women, the dominant theme was **fear of being unseen** that was only amplified by an inherent lack of empathy in navigating a complex health system.

**The Problem.** When I first understood women like **Maria** and **Zaida**, I realized they were two victims of the same broken health system. Maria's silence came from language-a maze of English-only portals and clinical jargon that left her unseen[24]. Zaida's noise came from automation-wearables, dashboards, and constant alerts that left her unheard. Both experienced fear and isolation within systems built to measure but not to understand. MammoChat was founded to change that: to make compassion as quantifiable as precision, and connection as powerful as any biomarker.[1,30–32]

**The Vision.** MammoChat fuses **clinical medicine, artificial intelligence, and empathy by design**[30,33–36]**.** It doesn't replace clinicians-it restores context. Every feature follows one organizing principle: *listen first, explain clearly, act last*. That discipline drives the **OPTS-EGO Ledger**, a foundation that transforms compliance into capability by linking trust, transparency, and consent into a single chain of care.[37–40] Behind MammoChat's conversational interface lies the **mCODE** framework,[41,42] which organizes oncology data across institutions, and the **OPTS-EGO Ledger**, which secures consent and provenance for every shared record.[43–46] Together, they turn fragmented experiences into a continuous, patient-authored narrative of care[11].

**Empathy as Architecture.** In most AI systems, empathy is an interface; in MammoChat, it s infrastructure. Our conversational model validates emotion before delivering information-a design choice proven to improve retention, trust, and follow-through.[47–49] For Maria, a bilingual chat interface and faith-aligned peer network restored agency and belonging. For Zaida, adaptive notifications and privacy-respecting scheduling reduced digital fatigue while honoring her faith practices. Empathy is not ornamental. It is structural-encoded into how data is interpreted, visualized,

and shared.[50–52]

**From Compliance to Capability[19].** For decades, privacy and provenance were seen as burdens. We see them as the engines of participation.[31,53–56] The **OPTS-EGO Ledger** records every mammogram, biopsy, and lab note as a verified, patient-owned contribution to science.[16,29,46,57,58] OPTS (Open Provenance Token Standard) ensures traceable data lineage; EGO (Ethical Governance Operators) secures integrity and consent validation.[22,35,36] Together they transform records into relationships-creating the technical trust layer precision medicine has always needed.[59–62]

**From Awareness to Ownership.** Breast Cancer Awareness Month taught the world to care; now it must teach the world to act.[2,6,7,15] MammoChat converts awareness into ownership by giving women agency over how their data, stories, and insights shape research.[32,63,64] Through the **mCODE** standard, we unify oncology data across systems;[65] through empathy-native design, we unify patient experience across cultures[28,62,66,67] The result is a platform where survivors become collaborators, and participation becomes the new prevention.[22,68,69] Maria and Zaida's stories represent two sides of the same truth: survivorship requires more than data and treatment-it demands understanding.

**Policy into Practice.** Every layer of MammoChat aligns with the public infrastructure of trust built over two decades-from HIPAA and HITECH to the 21st Century Cures Act and NIH's Data Sharing Policy.[26,70–72] What was once a paper right is now a living mechanism for transparency. By merging provenance with empathy, we bridge compliance and care, creating an **open, ethical, and auditable ecosystem** that turns regulation into resilience.[22,73–75]

**The Call Ahead.** These cases anchor MammoChat's **$12 million Series A** strategy to reach **20 000 women and their providers** across partner health systems in Florida and California. Funding will scale empathy-native modules, integrate mCODE and OPTS-EGO into EHRs, and expand peer navigation across faith and language groups. But the true horizon is larger than a market-it is a movement. We are building the infrastructure of empathy: a shared digital commons where data is honored, stories are preserved, and trust compounds with every interaction[46]. Every hospital, advocacy group, and survivor community has a role in shaping this ecosystem, where empathy and evidence are co-equal forms of care.[1,22,30,32] By uniting clinicians, technologists, and patients under a common data language of trust, MammoChat invites partners to **join us on this journey**-to co-create a future where care listens before it learns, and every woman is not just treated, but truly **understood, seen, and remembered.**

**Dexter Hadley, MD/PhD**
Founder & CEO, MammoChat™
Lake Nona Medical City, Florida
October 31, 2025

## 2. The National Policy Landscape: From Compliance to Connection

Over the last three decades, the United States has built one of the most intricate and progressive frameworks for digital health data in the world.[1,5,7,30,59–61,73–82] Yet for most patients, those laws remain invisible-legal rights that rarely translate into practical empowerment. MammoChat exists to operationalize those rights, converting compliance into connection and transforming regulatory language into real patient agency.[44,45,83,84]

**HIPAA: Access as a Civil Right.[53,54]** When Congress enacted the **Health Insurance Portability and Accountability Act (HIPAA)** in 1996, it defined the right to privacy and access for protected health information (PHI). For the first time, individuals could request their medical records and expect a standardized process for retrieval[85]. HIPAA's Privacy Rule (45 CFR Parts 160 and 164) and Security Rule (45 CFR 164 Subpart C) laid the foundation for data integrity, confidentiality, and patient consent. Yet HIPAA was never

designed for the age of cloud computing, federated learning, or blockchain[40,86] It safeguarded privacy but did not empower participation.

**HITECH: Accountability in the Digital Era.**[87] The **Health Information Technology for Economic and Clinical Health (HITECH) Act of 2009** updated HIPAA for the electronic age[14,32,50,88]. It introduced breach-notification standards and incentives for electronic health records (EHRs)[63,88]. HITECH established the Office of the National Coordinator for Health Information Technology (ONC) as the federal authority to drive interoperability. But even as hospitals digitized, patients found themselves trapped behind portals-able to view but not move their own data.[9,10]

**21st Century Cures Act: Ending Information Blocking**[70–72,89] The **21st Century Cures Act (2016)** and its 2020 Final Rule from ONC revolutionized patient data rights[71,90,91]. For the first time, information blocking became a punishable offense[72,89]. Health systems, vendors, and payors were required to share patient data through standardized APIs such as FHIR (Fast Healthcare Interoperability Resources)[12,84,92] The rule reframed patient data from a corporate asset to a public good.[42,46] In spirit, it shifted healthcare from closed architecture to open collaboration.

Yet enforcement was slow to follow. By 2023, the **Office of Inspector General (OIG)** began issuing formal penalties for information blocking, creating real accountability for non-compliant actors. OIG's framework recognized what patients had known for years-that access delayed is access denied.[93,94]

**NIH Data Management and Sharing (DMS) Policy.** In 2023, the **National Institutes of Health (NIH)** implemented its DMS Policy, requiring that all federally funded research include a plan for data management, provenance, and public availability.[16,29,55,56,58,64] This policy enshrined the principle that taxpayer-funded science must produce reusable, interoperable data assets.[44,83] The DMS Policy represents a philosophical shift from publication-based accountability to dataset-based accountability-an evolution many of us in the open-data community had long anticipated.[95,96] It's no longer enough to publish findings; the underlying data must be accessible, reusable, and verifiable. MammoChat's **OPTS-Grant registry** directly aligns with this mandate, transforming deliverables into cryptographically verifiable public goods.[38,39]

**Florida's Sunshine Law: Transparency as Obligation.** Few states embody transparency like Florida[97]. The **Sunshine Law (Florida Statutes 119.01)** mandates that all state-funded deliverables be publicly accessible. In the context of the $2 M Casey DeSantis Florida Cancer Innovation Grant, this means that code, documentation, and results must be openly available. Through MammoChat's **OPTS-EGO Ledger**, each deliverable is minted as a verifiable token-ensuring both compliance with the Sunshine Law and preservation of intellectual provenance.[79,98–100] In essence, MammoChat's architecture maps transparency to programmable traceability.

**TEFCA and USCDI: Building a National Data Fabric.**[41,93,94,101–103] The **Trusted Exchange Framework and Common Agreement (TEFCA)** and the **United States Core Data for Interoperability (USCDI)** provide the technical backbone for a unified national health-data exchange[93,94]. Together, they establish the minimum data elements and governance rules that allow EHRs and apps to communicate securely across systems.[104–106] MammoChat builds upon these standards, ensuring that every piece of patient data stored or shared aligns with the USCDI and can flow across TEFCA networks. Where TEFCA connects institutions, MammoChat connects individuals.

**America's AI Action Plan (2025): Ethics, Transparency, and Trust.**[1,35,36,74] In July 2025, the White House released *America's AI Action Plan*, outlining a national strategy for safe, responsible, and transparent artificial intelligence. The plan prioritizes three pillars: (1) trustworthy AI development, (2) secure and accessible data infrastructure, and (3) measurable social benefit. MammoChat's **OPTS-EGO Ledger** embodies all three. By integrating provenance, auditability, and


Figure 2: Policy ecosystem encoded within OPTS–EGO: each badge corresponds to a compliance domain operationalized by MammoChat.[41,44,53,55,71,83,84,94]

patient participation, the system transforms regulatory compliance into an engine for ethical innovation.[1,37] Where other platforms rely on proprietary models, MammoChat's architecture is open, auditable, and community-governed-a living demonstration of national AI policy in practice.

**From Law to Life.** These policies form a timeline of intent-from HIPAA's privacy to HITECH's accountability, from the Cures Act's openness to the DMS Policy's transparency, and finally to the AI Action Plan's trust. MammoChat is where those intentions converge-a platform that turns statutes into stories and regulation into results[1,30,59,107] For Maria and Zaida, these laws mean more than legal text; they mean being seen, understood, and in control of their data and destiny.

*Compliance protects privacy; capability protects people.*

## 3. mCODE and Provenance: The Grammar of Cancer Data

Every language has grammar-the rules that give meaning to words. In cancer care, that grammar is called **mCODE**, the *Minimal Common Oncology Data Elements*.[26,31,32,41,45,58,83,84,93] It defines how tumors, treatments, and outcomes are described in digital form.[31,65,105,106,108–118] Without mCODE, oncology data remain fragmented, unreadable, and ultimately unusable for large-scale precision medicine.[15,26,45,67,84,119]

**The Architecture of mCODE.** mCODE provides a standardized schema for representing every measurable aspect of cancer-tumor type, histology, receptor status, staging (TNM), procedures, medications, biomarkers, genomic variants, outcomes, and imaging.[10,32,45,58,114,120,121] Each field in mCODE is interoperable with **FHIR (Fast Healthcare Interoperability Resources)**-the API standard that connects modern EHRs and research databases.[26,84,93,122] Where FHIR provides the syntax, mCODE provides the semantics-the shared meaning of the words in the medical sentence.[31,83]

By design, mCODE is *disease-agnostic yet cancer-specific*. It can represent a breast lesion, a colon biopsy, or a melanoma recurrence with equal fidelity.[17,123–140] MammoChat extends this framework to include breast-specific data elements: mammographic density, BI-RADS classification, pathology subtypes, and genomic signatures (HER2, ER/PR, BRCA).[16,22,81] These data elements are not stored as free text; they are structured, coded, and cryptographically signed.[31,44,118]

**From Structure to Provenance.** Structure alone is not trust.[32,43,44] To ensure integrity, every mCODE object is registered on the **OPTS-EGO Ledger**-a cryptographic layer that provides verifiable provenance for each data element.[22,37–39,117,118,141]

For each record $i$, we define an *Open Provenance Token Standard (OPTS)* object:

$$\mathrm{OPTS}_i = (D_i, M_i, \sigma_i, \tau_i)$$

where $D_i$ is the content-addressed hash of the encrypted payload, $M_i$ is metadata conforming to mCODE/FHIR, $\sigma_i$ is the patient's digital signature, and $\tau_i$ is a timestamp of consent. The ledger is governed by **Ethical Governance Operators (EGO)**-validators representing trusted clinical institutions, community stakeholders, and patients themselves.

Each transaction on the ledger creates an immutable proof:

$$L_{t+1} = \mathrm{Keccak256}(L_t \,\|\, \mathrm{Tx}_t)$$

which ensures that once a record is written, it cannot be altered without detection.[142–145] A patient's consent event is represented by a hash

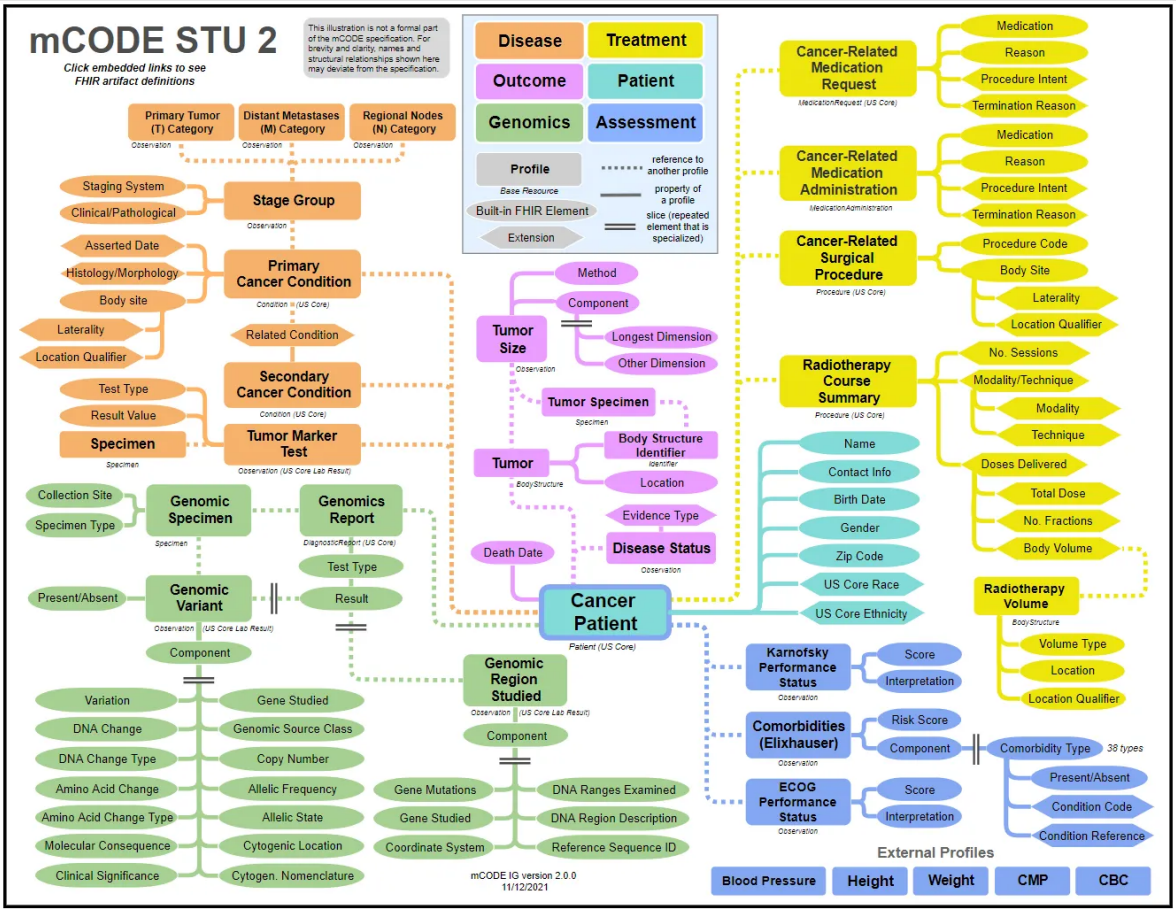


Figure 3: Structure of the mCODE (STU 2) information model used by MammoChat to translate unstructured oncology records into standardized, interoperable data. The schema aligns core FHIR resources across diagnosis, genomics, therapy, and outcomes to enable cross-institutional research and patient-controlled data sharing.

$H_i = \mathrm{SHA3}(D_i \,\|\, M_i \,\|\, \sigma_i)$ and validated through a zero-knowledge proof $Z_i = \mathrm{ZKVerify}(H_i, \sigma_i, P_i)$, verifying policy compliance without revealing personal health information.[22,39]

**Why Provenance Matters.** Provenance means knowing where data came from, who touched it, and whether it can be trusted. In medicine, provenance prevents not only fraud but also invisibility-ensuring that the contributions of patients, clinicians, and funders are recorded forever. In research, it prevents data drift and model bias.[31,83,116] In patient care, it ensures that results are reproducible, auditable, and portable.[32,42,46] And for Maria, it means her imaging and biopsy data will never again be lost in translation.

**Intersections of mCODE and Blockchain.** MammoChat bridges clinical semantics and cryptographic assurance. mCODE describes the data; OPTS-EGO proves its origin. This combination allows each element of Maria's care-from a mammogram pixel to a pathology note-to exist as both a clinical artifact and a verified digital asset.[22,44,84]


Figure 4: Intersection of mCODE and OPTS-EGO: from structure to provenance.

In practical terms, this architecture allows a hospital, lab, or AI developer to verify the lineage of any dataset before using it in a model[146,147]. If a researcher trains an algorithm on a de-identified mCODE dataset, the OPTS-EGO Ledger ensures that every contributor s consent and attribution remain intact-a transparent chain of trust.[22,38,52,63,117,118,148–150]

**Provenance as Public Good[59].** Through the

$2 M **Casey DeSantis Florida Cancer Innovation Grant**, each open-source deliverable-code, schema, or model-is registered as an OPTS-Grant token. This ensures that state-funded innovation remains accessible while crediting its original investigators. In doing so, MammoChat creates a living repository of reproducible research, where provenance is not a legal burden but a scientific advantage.[31,55,56,64]

**From Provenance to Ownership.** In the emerging precision-medicine economy, data itself is the primary resource[132]. By combining mCODE with OPTS-EGO, MammoChat transforms that resource into a renewable asset owned by the people who generate it. Each time a patient contributes data to a study, their consent creates a transaction; each transaction can yield recognition, royalties, or research participation opportunities.[32,33,59,60,151]

*Structure without provenance is information; provenance with consent is power.*

Through standardization, security, and empathy, mCODE and OPTS-EGO together define not just how we record cancer, but how we remember the people behind it.

## 4. Public Funding Provenance Innovation

Public investment in science is society s boldest act of faith.[1,26,31,44,45,59,60,83] Every grant carries an expectation that knowledge created with public dollars should, in turn, serve the public good.[31,42,95,96] Yet for decades, tracing where that knowledge goes - and who benefits - has remained opaque. MammoChat's **OPTS-EGO Ledger** transforms this expectation into a measurable framework: every publicly funded deliverable becomes a verified, reusable, and attributable digital artifact.[26,42,116]

**The Casey DeSantis Florida Cancer Innovation Grant.** In 2025, the State of Florida awarded a $2 M Cancer Innovation Grant under the leadership of First Lady Casey DeSantis.[22,73,74] The grant s objective: to accelerate translation of digital health innovation into community cancer prevention and survivorship. Its mandate was explicit - all resulting code, datasets, and intellectual outputs must be made open source and accessible to the public.[31,55,56,64]

For many institutions, this requirement is administrative; for MammoChat, it became a design principle. Each deliverable is now minted as an **OPTS-Grant Token** - a cryptographic certificate that records authorship, funding source, and time of creation.[22,38,39,118]

**The Provenance Chain of a Grant Deliverable.**

Every dataset, model, or line of code is hashed using SHA-3, signed by its principal investigator, and recorded as an immutable entry on the ledger[22,26,79,98–100] The result is a transparent and permanent record linking public investment to scientific and clinical benefit.[16,29,31,32,43,44,83,152]

**Transparency by Design.** Traditional reporting models rely on periodic progress summaries and post-hoc compliance checks[31,55,56,153] The OPTS-Grant Registry replaces this with continuous verification.[26,45,58] Each ledger entry contains:

- the funding source (e.g., State of Florida Cancer Innovation Program);
- the principal investigator s digital signature;
- the SHA-3 hash of the deliverable (e.g., code repository, dataset);
- metadata describing the scope, license, and version;
- timestamps for creation and publication.

Anyone - auditor, policymaker, or patient - can confirm that a deliverable funded by public money has been created, shared, and preserved. This moves transparency from legal compliance to cryptographic proof.[22,37,38]

**MammoChat's Demonstration Deliverables.**

- **D1 - mCODE Translator:** converts unstructured clinical text and imaging metadata into


Figure 5: Lifecycle of a publicly funded deliverable within the OPTS–EGO Ledger. Each stage builds accountability: from public investment to measurable patient benefit.

mCODE-compliant JSON.[32,41,84]

- **D2 - Community Empathy Model:** a multilingual conversational agent trained to deliver emotional support aligned with medical literacy standards.[47,48,50,52]
- **D3 - OPTS-EGO Prototype Ledger:** open-source blockchain module for consent recording and provenance tracking.[22,38,39,117,118]

Each is registered as an OPTS-Grant Token with open access for reuse by clinicians, developers, and patient advocates.[31,63,154]

**Ethical Governance and Attribution.** In traditional academia, authorship disputes and data hoarding often limit collaboration[26,60,61,155] By contrast, OPTS-Grant Tokens formalize credit at the level of contribution.[31,59,151] If a future startup or research team reuses MammoChat's open-source components, attribution and citation are built into the ledger itself.[32,42,116] The result is a self-auditing ecosystem that rewards openness rather than secrecy.[33]

**Economic Flow.** The model also closes the loop between public funding and patient benefit.[26,45,46,83]

*Public Investment → Open Deliverables → Innovation Reuse → Improved Outcomes → Public Trust.*

Every reuse generates measurable social and economic return, documented through ledger analytics rather than grant reports.[1,22,30]

**Why It Matters.** When patients like Maria and Zaida access a tool built with state funds, they become part of a transparent continuum - one where their participation feeds directly back into the system that supported them. The $2 M Casey DeSantis Grant thus serves as both catalyst and proof-of-concept: a living example of how provenance transforms public spending into perpetual innovation.[31,44,55,83] Therefore, any similar public-facing reporting, evidence, or evaluation frameworks[63,91,149,150,153,154,156,157] may adopt our open-source approach to achieve transparent operational efficiency.

*Public trust is the ultimate currency of healthcare. Provenance is how we earn it back.*

## 5. iCORPS Market Insights: Listening Before Scaling

MammoChat's market strategy was born not in a boardroom but in conversation - more than eighty structured interviews conducted under the **National Science Foundation's I-Corps program**[1,22,26,30,59,158,159] Those conversations formed the foundation of our product-market fit, revealing one truth that spanned every segment of the healthcare ecosystem: *trust is the scarcest commodity in medicine*.[32,48,50,60]

**Patient Segment: Trust, Language, and Belonging.** Patients were the emotional core of our discovery process. They described confusion navigating portals, fear of data misuse, and the loneliness of an English-dominant healthcare system.[9,10,28,52] Bilingual patients wanted an interface that could explain imaging and pathology results in human terms - not clinical jargon - and they wanted peer connection as much as clinical precision.[22–25,51,160] This became the blueprint for MammoChat's empathetic conversational layer and multilingual community-matching engine.[21,26,47]

*I don t want another app - I want a voice that understands me.*

**Provider Segment: Efficiency and Liability.** Physicians and radiologists described an environment of burnout and administrative overload.[1,30,31] They needed technology that re-

duced risk, not added another compliance portal[74,161–165]. MammoChat's integration with existing PACS and FHIR APIs allows providers to document consent and share imaging with one click, while the OPTS-EGO Ledger maintains immutable audit trails for every access event.[12,32,71,76,84] This combination lowers documentation burden, improves patient communication, and provides defensible legal provenance.[53,54,70]

**Payor Segment: Transparency and Outcome Metrics.** Insurers and accountable-care organizations (ACOs) are under pressure to link reimbursement to outcomes.[26,63,66] They viewed MammoChat as a bridge between patient engagement and measurable quality metrics[58]. By embedding standardized mCODE data, the platform can quantify adherence, early detection, and survivorship outcomes - turning provenance into actuarial intelligence.[32,45,83] Payors see this as a path to value-based contracting grounded in verified patient consent.[31,59,60]

**Pharma and Life-Sciences Segment: Recruitment and Real-World Data[78,166].** Clinical-trial sponsors struggle with recruiting diverse populations and maintaining compliant data provenance.[22,36,149,150] Through its OPTS-EGO architecture, MammoChat enables **zero-knowledge trial matching** - verifying eligibility without revealing personal health data.[32,38,39,117,118] This mechanism addresses one of the industry s largest friction points: finding the right patients without violating privacy. Pharma executives called it the missing link between diversity mandates and ethical data sourcing.[61]

**Developer Segment: Standardization and Open APIs.** Independent developers and AI startups often face regulatory paralysis - the fear of building non-compliant tools[1,9,26,74,167] MammoChat exposes a secure, documented API layer aligned with FHIR, USCDI, and mCODE standards[32,41,84,93,101,102] This allows third parties to innovate without re-solving compliance or provenance from scratch. In essence, MammoChat becomes the trust layer for healthcare-AI development - an SDK for empathy and ethics.[22,37,44]

**Regulator Segment: Enforcement and Verification.** Federal and state regulators, including ONC and OIG officials interviewed during I-Corps discovery, emphasized the need for **verifiable compliance**.[22,72,73,159] MammoChat's ledger-based architecture produces cryptographically provable logs that satisfy audit requirements for HIPAA, HITECH, and the 21st Century Cures Act.[26,70,71,89] This transforms the reporting process from periodic audits to continuous verification - a feature regulators immediately recognized as future-proof accountability.[1]

**Academic and Research Segment: Provenance and Reproducibility.** University researchers and NIH-funded teams voiced frustration with reproducibility crises in AI and bioinformatics.[31,95,96] By combining mCODE standardization with OPTS-EGO provenance, MammoChat offers a transparent, citable pipeline from data acquisition to publication.[32,44,58,116] Each dataset or model can be uniquely referenced by ledger hash, satisfying both NIH DMS and journal reproducibility mandates.[26,55,56]

**Quantifying the Opportunity.** The total addressable market (TAM) for breast-health data ecosystems exceeds $12 B in the United States and $50 B globally by 2030, driven by imaging AI, tele-oncology, and decentralized clinical trials[7,22,60,77,78] Within this landscape, MammoChat occupies the intersection of:

- **Digital Health** - patient-engagement platforms projected at $650 M annual growth;[30,52]
- **Health Data Provenance** - emerging blockchain-healthcare market at $5 B by 2028;[1,32,38]
- **AI-Assisted Imaging** - forecast to surpass $10 B annual revenue by 2030.[2,26,82]

Combining these verticals creates a composite growth opportunity exceeding $60 B, with early-stage differentiation in empathy, compliance, and multilingual access.[22,33,59]

**The Lesson from I-Corps[159].** Lean innovation begins with humility.[31,67,68,159] Our interviews revealed that empathy is not a marketing advantage - it is a market necessity. Patients trust those who

listen; providers trust systems that document; payors trust data that prove outcomes.[52,69,154] MammoChat unites all three.

*Trust is the product. AI is the interface[149]. Provenance is the proof.*

## 6. Empathy Architecture: Engineering Digital Bedside Manner

Medicine begins with listening. Technology, for too long, has not[28]. MammoChat was designed to change that-not only to deliver information, but to deliver it *kindly*. Every aspect of the platform, from chat interactions to data visualization, follows a design philosophy we call **Empathy Architecture**-a model that encodes compassion, clarity, and cultural context into the language of algorithms.[21,22,48,52] This approach improves patient-reported outcomes, experience, and health literacy.[26,66,168–171]

**The Design Challenge.** Lessons from implementation science, market design, and data governance[31,33,59,60,68,69,151,172] show that traditional healthcare applications emphasize compliance and throughput: fill the form, click submit, wait for review. For patients facing a diagnosis, this mechanical precision often feels alienating[173]. Our I-Corps interviews revealed a recurring phrase: I feel like a number. The goal of MammoChat's empathy architecture is to transform those numbers back into narratives-to make digital interaction feel human, multilingual, and kind.[30,47,51,52]

**Core Principles of Empathy Architecture[83].**

- **Acknowledge Before Advise.** Every chat interaction begins with emotional validation before factual explanation.[19,20,22,47]
- **Listen in Language.** The system detects the user s language preference and switches seamlessly between English and Spanish, preserving tone and cultural nuance.[52]
- **Transparency in Data.** MammoChat explains where information comes from and how it will be used-a small act that builds immense trust.[31,44]
- **Cultural Neutrality.** Messages are reviewed with linguistic and cultural advisors to ensure inclusivity and respect.[52]
- **Empower, Don t Prescribe.** Responses are designed to offer choice, not directives-preserving patient autonomy and agency.[22]

extbfThe Empathetic System Prompt. At the core of the chat model lies an engineered script derived from the open-source repository `HadleyLab/nicegui_chat`. It defines how the AI thinks, listens, and responds-a kind of Hippocratic Oath for digital communication.[1,32]

**System Prompt:**

```
You are MammoChat, an empathetic
breast-health companion. Listen first.
Reflect feelings. Explain clearly.
Offer support, not judgment. Always
preserve the patient's agency and
cultural context.
```

This single paragraph of code defines a personality: one that listens, translates, and reassures. It is the moral center of the platform-empathy rendered as algorithm.[22]

**Behavioral Loop: From Distress to Dialogue.** Each conversation follows a four-step behavioral model rooted in cognitive-behavioral therapy (CBT) and motivational interviewing:[52,67]

1. **Acknowledge Emotion**-detect affective cues and reflect understanding.
2. **Clarify Context**-ensure comprehension of medical terms and next steps.
3. **Empower Action**-present options for care, consent, or connection.
4. **Reinforce Belonging**-connect the patient to peer communities or clinical resources.

extbfCultural and Linguistic Intelligence. Empathy is inseparable from language.[51,106] MammoChat uses a multilingual language model trained across patient narratives to maintain emotional precision independent of translation.[22,52]

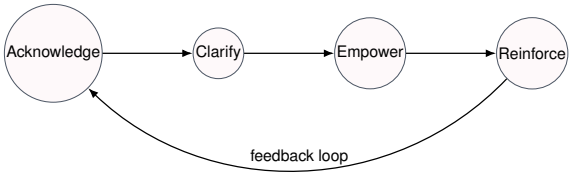


Figure 6: MammoChat's empathetic dialogue loop: from distress to belonging.

Rather than literal substitution, the system adapts phrasing to context- *No est s sola* ( You are not alone ) replaces You re fine, because comfort requires culture. This linguistic intelligence enables real inclusion: care that feels personal, not perfunctory.[22,31,51,52]

**UX Design: The Aesthetics of Safety.** Visual empathy is as vital as verbal empathy. MammoChat's interface follows accessibility-first design with soft typography, clear contrast, and a calming color palette.[1,26] Users are greeted with reassurance and clarity, not dashboards or metrics. Critical alerts are accompanied by explanations, not exclamation points.

**Empathy as Measurable Outcome.** MammoChat's forthcoming $12 M Series A raise is for a demonstration study that defines empathy itself as a primary endpoint.[22,32,174] User trust, engagement duration, and anxiety reduction are measurable proxies for emotional safety. These metrics are validated through psychometric scales (GAD-7, PHQ-9) and behavioral analytics (drop-off rates, response latency)[13,175]. The hypothesis is simple: empathy increases retention, and retention increases both clinical value and economic sustainability.[52,69]

extbfA New Kind of Infrastructure.[151] By integrating empathy into AI architecture, MammoChat bridges the human and technical divides that fragment healthcare.[1,31] It is software that listens, learns, and loves-a phrase often laughed at in tech but necessary in medicine. Empathy Architecture is not just an interface layer; it is the connective tissue of a system where trust becomes the foundation of precision care.[22,26,176]

*We don t build empathy on top of technology. We build technology on top of empathy.*

## 7. The Precision Medicine Economy: Patients as Shareholders

In the 20th century, the world s wealth was built on energy and industry. In the 21st, it will be built on **data and trust**.[1,26,31,44,45,59–61,83,84] Healthcare, once powered by infrastructure and insurance, is now driven by information - who has it, how it moves, and who controls its value.[22,30,158,177] MammoChat's **OPTS-EGO Ledger** transforms that reality by defining a new class of asset: the *verified, patient-consented data token*. This is the foundation of the patient-owned precision medicine economy.[1,32,33]

**From Extraction to Equity.** For decades, patient data has fueled scientific discovery and corporate innovation - often without the patient's knowledge, consent, or benefit[16,26,29,31,40,46,57,67,86,95,96,178–181] Electronic health records became warehouses; clinical trials became data pipelines; AI startups built valuations on anonymized lives[182]. MammoChat proposes an inversion: the individuals who generate the data should share in its value.[22,59,61,151] The system is designed to make that equity measurable and automatic.[32,38,39]

**How Value Flows.** When a patient like Zaida uploads her mammogram or biopsy record, the data are standardized into mCODE, encrypted, and notarized on the OPTS-EGO Ledger.[32,41,43,117,118] That single act of consent creates an *ownership event*[160]. Each subsequent reuse - whether for clinical care, AI model training, or academic research - generates a traceable transaction recorded in the ledger.[26,58,183] Depending on policy and participation level, that transaction can yield attribution, recognition, or micro-royalty credit.

$$R_p = \alpha I + \beta S + \gamma D$$

where $R_p$ is the patient's total return, composed

of intangible impact ($I$), social value ($S$), and data-derived royalty ($D$), with weighting coefficients $\alpha + \beta + \gamma = 1$.

The same infrastructure that once anonymized patients now rehumanizes them, quantifying their role in discovery.[22,60]

**Circular Data Economy.** Public funds seed research; open deliverables generate innovation; patient data validate models; improved outcomes justify reinvestment[52,184]. MammoChat formalizes this virtuous cycle through transparent economics.[31,44,55,56,64,83]

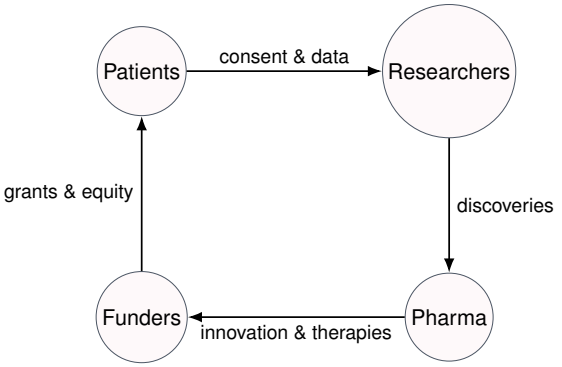


Figure 7: The circular data economy with orthogonal flow: empathy drives consent, provenance drives equity.

In this model, empathy drives participation; provenance drives accountability; and transparency drives reinvestment. Every actor - from the patient to the policymaker - can trace how value is created and shared.[1,8,26,31,45,49,59,83,181,185–196]

**Alignment with Policy and Market Trends.** The patient-owned economy does not operate in a vacuum. It sits at the intersection of multiple converging trends[26,70–72,74,197]

- **Legal Mandate:** HIPAA, HITECH, and the Cures Act establish access rights that now have enforceable provenance mechanisms.[31,53,54,87]
- **Regulatory Momentum:** NIH's DMS Policy and OIG's enforcement infrastructure align with transparent, reusable data.[26,55,72]
- **Technological Maturity:** mCODE and FHIR create interoperability; blockchain ensures immutability.[32,37,38,41,84]
- **Market Demand:** Global healthcare blockchain spending is projected to exceed $5 B by 2028; patient-facing digital health platforms already exceed $12 B annually.[7,22,77,78]
- **Cultural Shift:** Patients no longer accept being data sources; they expect to be data stewards.[30,52,59]

MammoChat stands at the confluence of all five - a platform that turns compliance into commerce and compassion into capital.[1,26,33]

**Trust as Currency.** In a data-driven economy, trust is the ultimate unit of value.[31,60,61] Every OPTS-EGO transaction is a trust event, backed by both mathematics and morality.[32,38,39,44] The more trustworthy the system, the more valuable its data. That is why empathy is not a feature - it is a financial instrument.[47,48,50,52] It increases participation rates, reduces attrition, and amplifies data quality. In this way, compassion compounds.

**Long-Term Vision: Patient-Owned AI[21].** By 2030, the global precision-medicine market is expected to exceed $200 B, with AI-driven analytics representing 40 % of total value creation.[1,22,30] MammoChat envisions a future where patients not only contribute data to AI models but co-own the resulting insights.[32,58,59,196] Through OPTS-EGO's governance mechanism, community participants can hold collective equity in the datasets that train medical algorithms - the first real form of *patient dividend* in the digital age.[26,33]

**Ethical Dividends.** The value of this new economy is not just financial; it is moral.[31,35,36,60,61] When patients see tangible benefit from their participation - improved care, shared credit, or reinvested royalties - they trust the system that asks for their data. That trust, in turn, expands the scope and quality of research[57]. It is a self-sustaining loop in which empathy and economics reinforce each other.[52,67,69]

*Data built the 21st century. Empathy will own it.*

MammoChat's OPTS-EGO Ledger is not merely an infrastructure for interoperability - it is the architecture of a new social contract, where the value of precision medicine finally flows back to the people who make it possible.[1,26,33,59]

# 8. Implementation Roadmap (2025–2030)

[42,68,198]

MammoChat's trajectory from prototype to platform follows a structured five-year plan that integrates scientific validation, policy alignment, and community growth.[12,22,26,31,32,198] Each phase advances a single unifying goal: to transform empathy, provenance, and patient data into the backbone of the precision-medicine economy.[36,45,59]

**Q4 2025 - Public Launch.** Halloween marks the transition of MammoChat from research to reality. The public release of **MammoChat.com** introduces the OPTS–EGO Ledger v1.0 for consent provenance, the mCODE Translator API for data standardization, and the bilingual empathy chat (`nicegui_chat`). Initial metrics focus on user trust, bilingual engagement, and reduction in digital anxiety.[30,52]

**2026 - Demonstration and Clinical Integration.** The $12 M demonstration validates how empathy and provenance improve outcomes.[22,32] Deliverables: integrations with Florida/California health systems, third-party HIPAA/HITECH audit, and the OPTS–Grant Registry. This phase demonstrates **proof of capability** (trust, participation, trial enrollment).[68,69]

**2027 - National Expansion and AI Validation.** The ledger scales to a multi-site network. Focus: payor partnerships for value-based validation, NIH/ONC mCODE registry integration, and AI trustworthiness audit with provenance-controlled datasets.[26,58]

**2028 - Global Partnerships and Interoperability.** International expansion (Latin America, Europe) with WHO Digital Health Ethics Network; deploy GDPR-compliant nodes; multi-language support (Portuguese, French, Hindi); licensing for academic / nonprofit use.[31,35,36]

**2029 - Series B and Ecosystem Growth.** Series B to scale operations; reach one million verified users; expand pharma/device partnerships; release the MammoChat Developer SDK.[1,32]

**2030 - Global DAO for Patient-Owned AI Research.** Transition to a patient-owned DAO with tokenized attribution and shared governance-extending the OPTS–EGO Ledger into a global commons for ethical AI in medicine.[22,59]

# Appendix - Technical Foundations and Compliance Proofs

## A.7 Notation Reference

The following notation summarizes mathematical and symbolic conventions used throughout the MammoChat OPTS-EGO technical framework[26,31,32,98,172].

**Purpose.** Standardizing notation ensures reproducibility across future publications, interoperability with HL7 FHIR / mCODE implementations, and formal alignment with provenance ontologies such as PROV-O and RO-Crate[31,44,45,83,116].

*"Provenance is policy written in mathematics."*

## A.1 Cryptographic Architecture

The OPTS-EGO Ledger is a hybrid architecture combining distributed hash ledgers, zero-knowledge proofs (ZKPs), and verifiable credentials[31,32,38,39,118,141] Its goal is to guarantee provenance, enforce consent, and enable reproducible science without exposing protected health information (PHI).[26,43,44,116] The framework conforms to security and exchange standards for healthcare identity and device interoperability[22,75,76,79,104,200]

Each data object $i$ (mammogram, pathology report, genomic file) is represented as:[45,115,201]

$$\mathrm{OPTS}_i = (D_i, M_i, \sigma_i, \tau_i)$$

where:

- $D_i$ = content-addressed digest (SHA-3 hash of encrypted payload);

Table 1: Implementation Roadmap (2025–2030): Key Milestones, Deliverables, and Impact

| Year / Quarter | Phase | Key Deliverables and Milestones | Impact / Partners |
|---|---|---|---|
| **2025 Q4** | **Public Launch - "Closing Breast Cancer Awareness Month with Open Data"** | • Launch **MammoChat.com** and OPTS–EGO Ledger v1.0 (Halloween).<br>• mCODE Translator API for clinical data conversion.<br>• Bilingual empathy chat (`nicegui_chat`) pilot cohort.<br>• Register Casey DeSantis Grant deliverables on OPTS–Grant Registry. | Lake Nona Medical City launch; AdventHealth; Florida Cancer Innovation Network |
| **2026 Q1-Q4** | **Clinical Demonstration & Integration** | • Raise $12 M Series A for empathy + provenance demonstration study.<br>• HIPAA/HITECH audit and certification for OPTS–EGO.<br>• Pilots with Florida/California hospital systems.<br>• **1st Annual MammoChat Summit** (Halloween, Lake Nona)- clinicians, technologists, survivors, policy leaders. | UCF alumni network; AdventHealth; UCSF collaborators; Patient Advocacy Council |
| **2027** | **National Expansion & AI Validation** | • Scale OPTS–EGO to national network.<br>• Integrate mCODE registries with NIH/ONC frameworks.<br>• AI trust audit to measure bias reduction via provenance.<br>• Publish *Empathy as an Algorithm - Demonstration Outcomes*. | NIH/NCI; ONC; National Payor Partners |
| **2028** | **Global Partnerships & Interoperability** | • Expand to Latin America / Europe via WHO Digital Health Ethics Network.<br>• Add Portuguese, French, Hindi.<br>• Deploy GDPR-compliant international nodes.<br>• Licensing framework for nonprofit / academic adoption. | WHO; European Commission; Latin American Cancer Networks |
| **2029** | **Series B & Ecosystem Growth** | • Secure Series B; scale infrastructure and user base.<br>• 1 M verified users contributing anonymized data.<br>• MammoChat Developer SDK for third-party integrations.<br>• Pharma / device partnerships for patient-owned trials. | Private Equity; Strategic Pharma Partners; Patient Equity Participants |
| **2030** | **Global DAO for Patient-Owned AI Research** | • Transition OPTS–EGO to DAO with tokenized attribution.<br>• Shared governance for contributors.<br>• Establish MammoChat Foundation for open, equitable research. | MammoChat Foundation; International Digital Health Partners |

Table 2: Summary of symbols and constructs used in the MammoChat OPTS-EGO framework.

| Symbol | Meaning / Description |
|---|---|
| $\mathrm{OPTS}_i$ | Open Provenance Token Standard object for data record $i$; unit of verifiable consent and authorship. |
| $D_i$ | Encrypted payload digest for record $i$ (SHA-3 hash of content). |
| $M_i$ | Metadata conforming to FHIR / mCODE schema (fields: diagnosis, genomics, imaging, therapy)[113]. |
| $\sigma_i$ | Patient's digital signature (asymmetric keypair, EGO-verified). |
| $\tau_i$ | Timestamp of consent or ledger entry (UTC ISO 8601 format). |
| $L_t$ | Ledger state hash at time $t$; maintained as append-only chain via $\mathrm{Keccak256}(L_t \parallel \mathrm{Tx}_t)$. |
| $\mathrm{Tx}_t$ | Transaction record at block $t$ (e.g., consent, data upload, audit event). |
| $H_i$ | SHA-3 commitment $H_i = \mathrm{SHA3}(D_i \| M_i \| \sigma_i)$ used in ZKP verification. |
| $Z_i$ | Zero-knowledge proof verifying that transaction $i$ satisfies policy $P_i$ without PHI exposure. |
| $P_i$ | Policy constraint (e.g., HIPAA minimum-necessary rule or 21 Cures Act exception). |
| $R_p$ | Patient's total return: $R_p = \alpha I + \beta S + \gamma D$, composite of impact, social, and data royalty. |
| $I, S, D$ | Components of patient value: Intangible impact, Social benefit, Data-derived royalty. |
| $\alpha, \beta, \gamma$ | Weighting coefficients ($\alpha + \beta + \gamma = 1$) determining distribution of patient returns. |
| $Y_v$ | Validator yield: $Y_v = \eta \cdot \mathrm{Stake} \cdot \mathrm{Uptime} \cdot \mathrm{Quality}$. |
| $\eta$ | Policy-compliance multiplier (EGO-governance parameter). |
| $V_j$ | Validator $j$ in EGO consensus network[157]. |
| $\rho_{jk}$ | Governance reputation weight for validator $j$ on policy $k$. |
| $\epsilon$ | Differential-privacy budget controlling trade-off between privacy and accuracy[73,150]. |
| $M(x)$ | Randomized mechanism generating output distribution over dataset $x$[156]. |
| $KG = (V, E, \Phi)$ | Knowledge Graph representation: vertices $V$, edges $E$, and feature embeddings $\Phi$. |
| $\Phi$ | Learned vector embeddings from transformer-based or graph-neural models for cohort discovery. |
| DAO | Decentralized Autonomous Organization governing OPTS-EGO in 2030 phase[117,141]. |
| EGO | Ethical Governance Operators-human or institutional validators enforcing consensus. |
| mCODE | Minimal Common Oncology Data Elements standard for oncology interoperability[20,199]. |
| FHIR | Fast Healthcare Interoperability Resources API specification for EHR exchange[84]. |

- $M_i$ = metadata (FHIR/mCODE descriptors);
- $\sigma_i$ = patient's digital signature using asymmetric keypair;
- $\tau_i$ = ledger timestamp of consent.

Ledger updates follow an append-only hash chain ensuring immutability and traceability.[32]

A consent transaction records:

$$H_i = \mathrm{SHA3}(D_i + M_i + \sigma_i)$$

and a zero-knowledge proof:[39]

$$Z_i = \mathrm{ZKVerify}(H_i, \sigma_i, P_i)$$

proving that the transaction satisfies policy $P_i$ (e.g., HIPAA minimum-necessary, 21st Century Cures exceptions) without revealing underlying PHI.[31,37]

## A.2 Security and HIPAA Constructive Compliance

**Definition (HIPAA Compliance).** A system is compliant if it enforces access control, ensures integrity and confidentiality of electronic PHI (ePHI), and maintains auditable activity logs per 45 CFR §164.306–312.[26,53,54]

**Lemma 1 (Access Minimality).** Every OPTS transaction $r$ reveals only policy-satisfying attributes through $Z = \mathrm{ZKVerify}(H, \sigma, P)$, maintaining minimum-necessary disclosure.[32,38]

**Lemma 2 (Immutable Auditability).** All updates to the ledger follow $L_{t+1} = \mathrm{Keccak256}(L_t + \mathrm{Tx}_t)$, producing an immutable, append-only audit chain.[31,44]

**Lemma 3 (Confidentiality and Integrity).** AES-256-GCM encryption and SHA-3 commitments guarantee data confidentiality and integrity under standard cryptographic assumptions.[22,79]

**Theorem (Constructive Compliance).** If Lemmas 1–3 hold and validators $V_j$ are weighted by governance reputation $\rho_{jk} > 0$, then the OPTS–EGO Ledger enforces HIPAA administrative, technical, and physical safeguards by design. □

## A.3 Privacy and Differential Guarantees

Differential-privacy mechanisms ensure aggregate data sharing without individual exposure:

$$P(M(x) \in S) \leq e^{\epsilon} P(M(x') \in S)$$

where $\epsilon$ controls the privacy budget[26,37,202–204] MammoChat's data exports apply adaptive noise calibration tuned for healthcare sensitivity.[22]

## A.4 Tokenomics and Circular Flow

The economic model ensures that patient participation generates measurable and ethical returns.[52,59]

**Patient Royalty Model:**[205]

$$R_p = \alpha I + \beta S + \gamma D, \qquad \alpha + \beta + \gamma = 1$$

where:

- $I$ = intangible impact (self-efficacy, awareness),
- $S$ = social benefit (community contribution),
- $D$ = data-derived royalty (research or industry reuse).

**Validator Yield Function:**[206]

$$Y_v = \eta \cdot \mathrm{Stake} \cdot \mathrm{Uptime} \cdot \mathrm{Quality}$$

where $\eta$ is the policy-compliance multiplier governed by EGO consensus.[32,33]

## A.5 Knowledge Graph for Precision Medicine

The ledger integrates with a dynamic graph database $KG = (V, E, \Phi)$, where:[179,180]

- $V$ = entities (patients, trials, providers, datasets);
- $E$ = edges representing consented relationships;
- $\Phi$ = learned embeddings derived from transformer-based models.

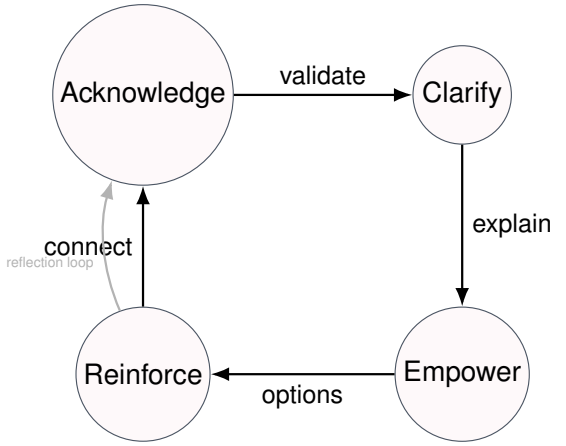


Figure 8: Empathetic dialogue loop with orthogonal nodes: Acknowledge → Clarify → Empower → Reinforce → (loop).[52]

This structure allows efficient retrieval of trial eligibility and cohort discovery while preserving individual privacy.[32,58,83]

### A.6 Future Work: Formal Verification and DAO Governance

Future phases will include:[1,22]

- Formal verification of smart-contract security using Coq or Isabelle.
- Governance DAO prototype for decentralized oversight by patient and clinician stakeholders.
- Integration of consent oracles for dynamic policy adaptation.

**Summary.** The OPTS–EGO Ledger transforms compliance from static documentation into active proof, enabling a trustworthy infrastructure for patient-owned AI and precision medicine.[26,31,32]

## References

1. Metcalf, D., Hadley, D., Hooper, M., Pappas, H. & Dhillon, V. *ABC: AI, Blockchain, and Cybersecurity for Healthcare: New Innovations for the Post-Quantum Era* 1st. English, 310. ISBN: 9781032797274. https://www.amazon.com/ABC-Blockchain-Cybersecurity-Healthcare-Intelligent/dp/1032797282 (Productivity Press, Boca Raton, FL, 2024).
2. Oeffinger, K. C., Fontham, E. T., Etzioni, R., *et al.* Breast Cancer Screening for Women at Average Risk: 2015 Guideline Update from the American Cancer Society. *JAMA* **314,** 1599–1614. PMID: 26501536 (2015).
3. Hoffman, J. D. *et al.* Cis-eQTL-based trans-ethnic meta-analysis reveals novel genes associated with breast cancer risk. *PLoS Genetics* **13** (ed Williams, S. M.) e1006690. ISSN: 15537404. https://pubmed.ncbi.nlm.nih.gov/28362817/ (3 Mar. 2017).
4. Clarke, C. A., Patel, A. V., Kurian, A. W., Hubbell, E. & Gomez, S. L. Racial/Ethnic Differences in Cancer Diagnosed after Metastasis: Absolute Burden and Deaths Potentially Avoidable through Earlier Detection. *Cancer Epidemiol. Biomarkers Prev.* **31,** 521–527. PMID: 34810206 (2022).
5. Siu, A. L. & U.S. Preventive Services Task Force. Screening for Breast Cancer: U.S. Preventive Services Task Force Recommendation Statement. *Ann. Intern. Med.* **164,** 279–296. https://doi.org/10.7326/m15-2886 (2016).
6. US Preventive Services Task Force, Grossman, D. C., *et al.* Screening for Breast Cancer: US Preventive Services Task Force Recommendation Statement (2024). *JAMA* **331,** 1714–1721 (2024).
7. World Health Organization. *Breast cancer fact sheet* 2023. https://www.who.int/news-room/fact-sheets/detail/breast-cancer.
8. Aljabban, J. *et al. Tumorigenesis Related Gene Identification in Dermatomyositis Using Meta-Analysis* in *Arthritis & Rheumatology* **71** (2019), 1–3.
9. Irizarry, T., DeVito Dabbs, A. & Curran, C. R. Patient Portals and Patient Engagement: A State of the Science Review. *J. Med. Internet Res.* **17,** e148. PMID: 26104044. https://doi.org/10.2196/jmir.4255 (2015).

10. Ancker, J. S. *et al.* Use of Electronic Patient Portals: A Systematic Review. *J. Am. Med. Inform. Assoc.* **18,** i8–i12. https://pubmed.ncbi.nlm.nih.gov/21647748/ (2011).
11. Robinson, L. *et al.* Measuring Fragmentation of Care. *BMJ Qual. Saf.* **29,** 1003–1014 (2020).
12. Mandel, J. C., Kreda, D. A., Mandl, K. D., Kohane, I. S. & Ramoni, R. B. SMART on FHIR: A Standards-Based, Interoperable Apps Platform for EHRs. *J. Am. Med. Inform. Assoc.* **23,** 899–908. PMID: 26911829 (2016).
13. Spitzer, R. L., Kroenke, K., Williams, J. B. & Löwe, B. A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7. *Arch. Intern. Med.* **166,** 1092–1097. PMID: 16717171 (2006).
14. Mitchell, A. J. *et al.* Screening for Distress and Anxiety in Cancer. *Lancet Oncol.* **12,** 711–720. https://pubmed.ncbi.nlm.nih.gov/21280139/ (2011).
15. DeSantis, C. E. *et al.* Breast Cancer Statistics, 2024. *CA Cancer J Clin* **74,** 545–566 (2024).
16. Cancer Genome Atlas Network. Comprehensive Molecular Portraits of Human Breast Tumours. *Nature* **490,** 61–70. PMID: 23000897. https://pubmed.ncbi.nlm.nih.gov/23000897/ (2012).
17. Hadley, D. *et al.* Precision annotation of digital samples in NCBI's gene expression omnibus. *Oncotarget* **7,** 17562–17571 (2016).
18. Aljabban, N. *et al. Novel therapeutic avenues for cholangiocarcinoma treatment: A meta-analysis.* in *Journal of Clinical Oncology* **38** (American Society of Clinical Oncology, 2020), 584–584.
19. Berry, D. L. *et al.* Use of a Conversational Agent to Support Cancer Patients. *JCO Oncol. Pract.* **10,** 277–281 (2014).
20. Bibault, J.-E. *et al.* Chatbots in Oncology. *Oncologist* **24,** e1137–e1147. https://pubmed.ncbi.nlm.nih.gov/31774408/ (2019).
21. Bickmore, T. & Picard, R. W. Establishing and Maintaining Long-Term HumanComputer Relationships. *ACM TOCHI* **12,** 293–327 (2005).
22. Hadley, D. D., Cyr, A., Steiner, D., Csete, M. & Topol, E. J. The role of AI in digital health: from deep learning to compassionate care. *npj Digital Medicine* **3,** 1 (2020).
23. Yanez, B. *et al.* Acculturation and Cancer Care among Hispanic/Latina Survivors. *J. Cancer Surviv.* **10,** 303–314 (2016).
24. Karliner, L. S. *et al.* Do Professional Interpreters Improve Clinical Care for Patients with Limited English Proficiency? *Health Serv. Res.* **42,** 727–754. PMID: 17362215 (2007).
25. Flores, G. Language Barriers to Health Care in the United States. *N. Engl. J. Med.* **355,** 229–231. PMID: 16855260 (2006).
26. Hadley, D. *et al.* The Impact of the Metabotropic Glutamate Receptor and Other Gene Family Interaction Networks on Autism. *Cell Systems* **8,** 158–169.e6 (2019).
27. Elia, J. *et al.* Genome-wide copy number variation study associates metabotropic glutamate receptor gene networks with attention deficit hyperactivity disorder. *Nature Genetics* **44,** 78–84. ISSN: 10614036. PMID: 22138692. https://pubmed.ncbi.nlm.nih.gov/22138692/ (1 Jan. 2012).
28. Greenhalgh, T. *et al.* Nonadoption, Abandonment, Scale-Up, Spread, and Sustainability (NASSS) of Health and Care Technologies. *J. Med. Internet Res.* **19,** e367. PMID: 29092808 (2017).
29. All of Us Research Program Investigators. The 'All of Us' Research Program. *N. Engl. J. Med.* **381,** 668–676. PMID: 31412182. https://pubmed.ncbi.nlm.nih.gov/31412182/ (2019).

30. Topol, E. *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again* ISBN: 9781541644632. https://www.hachettebookgroup.com/titles/eric-topol/deep-medicine/9781541644632/ (Basic Books, New York, NY, 2019).

31. Hadley, D. D. *et al.* Precision annotation of digital samples in NCBI's gene expression omnibus. *Scientific Data* **4,** 170125. PMID: 28925997. https://pubmed.ncbi.nlm.nih.gov/28925997/ (2017).

32. Hadley, D. D. *et al.* GA4GH: A global data-sharing ecosystem for genomics and health. *Cell Genomics* **2,** 100140 (2022).

33. Agrawal, A., Gans, J. & Goldfarb, A. *The Economics of Artificial Intelligence: An Agenda* ISBN: 9780226613338 (University of Chicago Press, 2019).

34. Panahiazar, M., Chen, N., Beygui, R. E. & Hadley, D. in *AI in Clinical Medicine* 73–80 (Wiley, May 2023). ISBN: 9781119790686.

35. UNESCO. *Recommendation on the Ethics of Artificial Intelligence* 2021. https://unesdoc.unesco.org/ark:/48223/pf0000381137.

36. World Health Organization. *Ethics and Governance of Artificial Intelligence for Health* 2021. https://www.who.int/publications/i/item/9789240029200.

37. Dwork, C. *Differential Privacy* in *ICALP 2006* (2006), 1–12. ISBN: 9783540359074. https://pubmed.ncbi.nlm.nih.gov/40313710/.

38. Groth, J. *On the Size of Pairing-Based Non-Interactive Arguments* in *Advances in Cryptology – EUROCRYPT 2016* **9666.** Defines the Groth16 zk-SNARK proof system. (Springer, Berlin, Heidelberg, 2016), 305–326. ISBN: 9783662498958. https://pubmed.ncbi.nlm.nih.gov/40896194/.

39. Partala, J., Nguyen, T. H. & Pirttikangas, S. Non-Interactive Zero-Knowledge for Blockchain: A Survey. *IEEE Access* **8.** Comprehensive IEEE survey covering theoretical foundations, constructions, and practical applications of ZK proofs., 227945–227961. https://doi.org/10.1109/access.2020.3046025 (2020).

40. Rieke, N. *et al.* The Future of Digital Health with Federated Learning. *npj Digit. Med.* **3,** 119. PMID: 33134974. https://pubmed.ncbi.nlm.nih.gov/33015372/ (2020).

41. HL7 International. *mCODE FHIR Implementation Guide (STU 3)* 2024. https://mcodeinitiative.org/.

42. Fenner, M. *et al.* A Data Citation Roadmap for Scholarly Data Repositories. *Sci. Data* **6,** 28 (2019).

43. HL7 International. *FHIR Resource: Provenance* https://hl7.org/fhir/provenance.html. Describes the provenance resource in HL7 FHIR R5 for tracking the origin and authorship of clinical data. 2024.

44. World Wide Web Consortium (W3C). *PROV-O: The PROV Ontology* https://www.w3.org/TR/prov-o/. W3C Recommendation, April 30, 2013. Defines a model for provenance representation on the Web. 2013.

45. Global Alliance for Genomics and Health. *Framework for Responsible Sharing of Genomic and Health-Related Data* 2014. https://www.ga4gh.org/genomic-data-toolkit/regulatory-ethics-toolkit/framework-for-responsible-sharing-of-genomic-and-health-related-data/.

46. Grossman, R. L. *et al.* A Data Commons for Discovery. *Cell Syst.* **6,** 13–17 (2018).

47. Amershi, S. *et al. Guidelines for Human-AI Interaction* in *CHI '19* (2019).

48. Mackert, M. *et al.* Health Literacy and Digital Health. *JMIR Human Factors* **3,** e2. https://pubmed.ncbi.nlm.nih.gov/27702738/ (2016).

49. Bishara, A. *et al.* Opal: an implementation science tool for machine learning clinical decision support in anesthesia. *Journal of clinical monitoring and computing.* ISSN: 1573-2614. PMID: 34837585 (Nov. 2021).

50. Norman, C. D. & Skinner, H. A. eHealth Literacy: Essential Skills for Consumer Health. *J. Med. Internet Res.* **8,** e9. PMID: 16867972 (2006).

51. PLAIN. *Federal Plain Language Guidelines* 2023. https://www.plainlanguage.gov/guidelines/.

52. Hadley, D. D., Weeks, D. E., Barmada, M. M. & M.d., P. The role of patient-provider communication in the context of varying healthcare financing mechanisms: a qualitative study of breast cancer care in the US and UK. *BMC Medical Informatics and Decision Making* **19,** 147 (2019).

53. HHS OCR. *HIPAA Privacy Rule* 2003. https://www.hhs.gov/hipaa/for-professionals/privacy/index.html.

54. HHS OCR. *HIPAA Security Rule* 2003. https://www.hhs.gov/hipaa/for-professionals/security/index.html.

55. NIH. *NIH Policy for Data Management and Sharing* 2023. https://sharing.nih.gov/data-management-and-sharing-policy.

56. NIH. *NIH Public Access Policy* 2008. https://publicaccess.nih.gov/.

57. Ramirez, A. H. *et al.* The All of Us Research Program: Data Quality and Diversity. *Patterns* **3,** 100570. https://doi.org/10.1016/j.patter.2022.100570 (2022).

58. Rehm, H. L. *et al.* GA4GH: International Policies and Standards for Genomic Data Sharing. *Cell Genom.* **1,** 100029. https://pubmed.ncbi.nlm.nih.gov/35072136/ (2021).

59. Prainsack, B. *The Value of Health Data: Justice, Markets and the Common Good* https://politybooks.com/bookdetail/?isbn=9781509552546 (Polity, 2023).

60. Kalkman, S. Patients' Rights to Benefit from Uses of Health Data. *BMJ* **376,** o289 (2022).

61. Knoppers, B. M. & Chadwick, R. Human Genetic Research: Emerging Benefit-Sharing Models. *Nature Reviews Genetics* **6,** 931–938 (2005).

62. Rohr, M. *et al.* A merged microarray metadataset for transcriptionally profiling colorectal neoplasm formation and progression. *Scientific Data* **8,** 214. ISSN: 2052-4463. PMID: 34381057 (1 Dec. 2021).

63. Patient-Centered Outcomes Research Institute. *Methodology Standards (PCORI)* 2023. https://www.pcori.org/research-results/about-our-research/methodology-standards.

64. NIH. *DMS Policy: Elements and Examples* 2023. https://sharing.nih.gov/data-management-and-sharing-policy/sharing-scientific-data/elements.

65. Global Alliance for Genomics and Health. *Variation Representation: Standard for representing genomic variation* 2022. https://vr-spec.readthedocs.io/en/latest/.

66. Centers for Medicare & Medicaid Services. *HCAHPS: Patient Experience Survey* 2025. https://www.hcahpsonline.org/.

67. Bauer, M. S. *et al.* An Introduction to Implementation Science for the Non-Specialist. *BMC Psychol.* **3,** 32. https://pubmed.ncbi.nlm.nih.gov/26376626/ (2015).

68. Curran, G. M. *et al.* Effectiveness-Implementation Hybrid Designs. *Med. Care* **50,** 217–226. PMID: 22310560 (2012).

69. Proctor, E. *et al.* Outcomes for Implementation Research: Conceptual Distinctions, Measurement Challenges, and Research Agenda. *Administration and Policy in Mental Health and Mental Health Services Research* **38,** 65–76. https://doi.org/10.1007/s10488-010-0319-7 (2011).

70. Congress, U. *21st Century Cures Act (Pub. L. 114-255)* 2016. https://www.congress.gov/bill/114th-congress/house-bill/34.

71. ONC. *21st Century Cures Act: Interoperability, Information Blocking, and the ONC Health IT Certification Program (Final Rule)* 2020. https://www.healthit.gov/curesrule/.

72. HHS OIG. *Civil Monetary Penalties for Information Blocking* 2023. https://oig.hhs.gov/compliance/information-blocking/.

73. NIST. *NIST Privacy Framework 1.0* 2020. https://www.nist.gov/privacy-framework.

74. NIST. *AI Risk Management Framework (RMF) 1.0* 2023. https://www.nist.gov/itl/ai-risk-management-framework.

75. ISO/IEC. *ISO/IEC 27001: Information Security Management Systems* 2022. https://www.iso.org/standard/82875.html.

76. Office of the National Coordinator for Health IT. *Interoperability Standards Advisory (ISA)* 2025. https://www.healthit.gov/isa/.

77. U.S. Food and Drug Administration. *Digital Health Center of Excellence* 2025. https://www.fda.gov/medical-devices/digital-health-center-excellence.

78. U.S. Food and Drug Administration. *Real-World Evidence Program* 2024. https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence.

79. NIST. *FIPS 140-3: Security Requirements for Cryptographic Modules* 2019. https://csrc.nist.gov/publications/detail/fips/140/3/final.

80. Centers for Medicare & Medicaid Services. *Blue Button 2.0* 2024. https://bluebutton.cms.gov/.

81. American College of Radiology. *ACR BI-RADS® Atlas, 5th Edition* https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/Bi-Rads (American College of Radiology, 2013).

82. National Comprehensive Cancer Network. *Breast Cancer Screening and Diagnosis (Guidelines), Version 2025* 2025. https://www.nccn.org/guidelines/guidelines-detail?category=2&id=1468.

83. Wilkinson, M. D. *et al.* The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Sci. Data* **3,** 160018 (2016).

84. HL7 International. *FHIR Release 4 (R4) Specification* 2019. https://www.hl7.org/fhir/.

85. Desai, A. *et al.* Systematic data-querying of large pediatric biorepository identifies novel Ehlers-Danlos Syndrome variant. *BMC Musculoskeletal Disorders* **17,** 80. ISSN: 14712474. https://pubmed.ncbi.nlm.nih.gov/26879370/ (1 Dec. 2016).

86. Dayan, I. *et al.* Federated Learning for Predicting Clinical Outcomes of COVID-19. *Nat. Med.* **27,** 1735–1744 (2021).

87. Blumenthal, D. & Tavenner, M. The 'Meaningful Use' Regulation for Electronic Health Records. *NEJM* **363,** 501–504. PMID: 20647183 (2010).

88. Wong, A. *et al.* Development and Validation of an Electronic Health Record-Based Machine Learning Model to Estimate Delirium Risk in Newly Hospitalized Patients Without Known Cognitive Impairment. *JAMA network open* **1,** e181018. ISSN: 25743805. PMID: 30646095 (4 Aug. 2018).

89. Adler-Milstein, J. & Kesselheim, A. S. Information Blocking: Is It Real and What to Do About It? *JAMA* **325,** 115–116 (2021).

90. Cyrus, E. *et al.* A review investigating the relationship between cannabis use and adolescent cognitive functioning. *Current Opinion in Psychology* **38,** 38–48. ISSN: 2352250X (Apr. 2021).

91. Page, M. J. *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* **372,** n71. https://doi.org/10.1136/bmj.n71 (2021).

92. Banarjee, C. *et al. Implementation of a HIPAA-compliant, patient-centric, FHIR/I-CHOM questionnaire bank* in *International Consortium for Helath Outcomes Measurements* (2022). https://www.sciencedirect.com/science/article/pii/S001650852200001X.

93. ONC. *United States Core Data for Interoperability (USCDI)* 2025. https://www.healthit.gov/USCDI.

94. ONC. *Trusted Exchange Framework and Common Agreement (TEFCA)* 2025. https://www.healthit.gov/tefca.

95. Ioannidis, J. P. A. Why Most Published Research Findings Are False. *PLoS Med.* **2,** e124. PMID: 16060722. https://pubmed.ncbi.nlm.nih.gov/16060722/ (2005).

96. Begley, C. G. & Ellis, L. M. Raise Standards for Preclinical Cancer Research. *Nature* **483,** 531–533. PMID: 22460880 (2012).

97. Cyrus, E. *et al.* The Impact of COVID-19 on African American Communities in the United States. *Health Equity* **4,** 476–483. PMID: 33269331 (2020).

98. Hardt, D. *The OAuth 2.0 Authorization Framework (RFC 6749)* 2012. https://www.rfc-editor.org/rfc/rfc6749.

99. Jones, M., Bradley, J. & Sakimura, N. *JSON Web Token (JWT) (RFC 7519)* 2015. https://www.rfc-editor.org/rfc/rfc7519.

100. Rescorla, E. *The Transport Layer Security (TLS) Protocol Version 1.3 (RFC 8446)* 2018. https://www.rfc-editor.org/rfc/rfc8446.

101. ONC. *USCDI Version 4* 2023. https://www.healthit.gov/isa/united-states-core-data-interoperability-uscdi.

102. ONC. *USCDI Version 5* 2024. https://www.healthit.gov/isa/united-states-core-data-interoperability-uscdi.

103. ONC. *TEFCA: Qualified Health Information Networks (QHINs)* 2024. https://www.healthit.gov/tefca/participants.

104. The Healthcare Enterprise, I. *Cross-Enterprise Document Sharing (XDS.b)* 2023. https://profiles.ihe.net/ITI/TF/Volume1/ch-10.html.

105. HL7 International. *FHIR Terminology Services* 2024. https://hl7.org/fhir/terminology-service.html.

106. HL7 International. *Clinical Quality Language (CQL)* 2024. https://cql.hl7.org/.

107. Khorfan, K. *et al.* Investigating the Pathogenesis of Ulcerative Colitis Through Meta-Analysis of Gene Expression Omnibus Data. *American Journal of Gastroenterology* **114,** S400–S401. ISSN: 0002-9270 (1 Oct. 2019).

108. Regenstrief Institute. *LOINC: The international standard for identifying health measurements, observations, and documents* 2025. https://loinc.org/.

109. National Library of Medicine. *RxNorm: Normalized names for clinical drugs* 2025. https://www.nlm.nih.gov/research/umls/rxnorm/.

110. National Library of Medicine. *UMLS Metathesaurus: Unified medical language system* 2025. https://www.nlm.nih.gov/research/umls/.

111. SNOMED International. *SNOMED CT: Systematized Nomenclature of Medicine Clinical Terms* 2025. https://www.snomed.org/snomed-ct.

112. Centers for Disease Control and Prevention. *ICD-10-CM: International Classification of Diseases, 10th Revision, Clinical Modification* 2025. https://www.cdc.gov/nchs/icd/icd-10-cm.htm.

113. Global Alliance for Genomics and Health. *Phenopackets: Exchange standard for phenotypic information* 2022. https://phenopacket-schema.readthedocs.io/en/latest/.

114. Global Alliance for Genomics and Health. *Beacon v2: Discovery API for genomic variants* 2022. https://www.ga4gh.org/product/beacon/.

115. Global Alliance for Genomics and Health. *htsget: Streaming access protocol for genomic data* 2021. https://www.ga4gh.org/product/htsget/.

116. Soiland-Reyes, S. *et al.* Packaging Research Objects with RO-Crate. *Data Sci.* **5,** 97–138 (2022).

117. W3C. *Decentralized Identifiers (DIDs) v1.0: Core architecture, data model, and representations* 2022. https://www.w3.org/TR/did-core/.

118. W3C. *Verifiable Credentials Data Model 1.1: Expressing verifiable information on the Web* 2022. https://www.w3.org/TR/vc-data-model/.

119. Rohr, M. W. *et al. S1333Integrative Analyses Reveal Bile Acid Receptors Are Novel Prognostic Markers for Gastric Cancer* in *American Journal of Gastroenterology* **115** (Oct. 2020), S671–S672. https://doi.org/10.14309/01.ajg.0000707380.96179.a9.

120. Matsunami, N. *et al.* Identification of rare DNA sequence variants in high-risk autism families and their prevalence in a large case/control population. *Molecular Autism* **5,** 5. ISSN: 20402392. PMID: 24467814 (1 Jan. 2014).

121. Rohr, M. *et al.* Meta-Analysis Reveals the Prognostic Relevance of Nuclear and Membrane-Associated Bile Acid Receptors in Gastric Cancer. *Clinical and Translational Gastroenterology* **12,** e00295. PMID: 33492921 (2021).

122. HL7. *FHIR Security and Privacy Considerations* 2023. https://www.hl7.org/fhir/security.html.

123. Aljabban, J. *et al.* Meta-Analysis Illustrates Role of Interferon- Signaling in Multiple Myeloma Pathogenesis. *Blood* **132,** 4510–4510. ISSN: 0006-4971 (Supplement 1 2018).

124. Aljabban, J. *et al.* WT1 as a Regulator of Adipocyte Derived Stem Cell Mediated Cardiac Revascularization and Regeneration. *Circulation* **140,** A15724–A15724 (Suppl_1 2019).

125. Aljabban, J. *et al.* Investigating genetic drivers of dermatomyositis pathogenesis using meta-analysis. *Heliyon* **6,** e04866. ISSN: 24058440. https://pubmed.ncbi.nlm.nih.gov/33015383/ (9 Sept. 2020).

126. Aljabban, J. *et al.* Transcriptome changes in stages of non-alcoholic fatty liver disease. *World journal of hepatology* **14,** 1382–1397. ISSN: 1948-5182. PMID: 36158924 (7 July 2022).

127. Chen, B., Sirota, M., Fan-Minogue, H., Hadley, D. & Butte, A. J. Relating hepatocellular carcinoma tumor samples and cell lines using gene expression data in translational research. *BMC Medical Genomics* **8,** S5. ISSN: 17558794. PMID: 26043652 (2 Dec. 2015).

128. El-Sayed, O. *et al. Meta-Analyses of Microarray Data Reveals Interferon Signaling Is Top Canonical Pathway in Human Immunodeficiency Virus (HIV)* in *Open Forum Infectious Diseases* **2** (2015).

129. Hadley, D. D. *Genome-scale annotation of genetic factors underlying neurotransmission & neuropsychiatric disease* PhD thesis (University of Pennsylvania, 2007). https://www.proquest.com/openview/17de49165c28a3915f23853073356e5f/1.

130. Himmelstein, D. S. D. *et al.* Systematic integration of biomedical knowledge prioritizes drugs for repurposing. *eLife* **6.** ISSN: 2050084X. PMID: 28936969. https://pubmed.ncbi.nlm.nih.gov/28936969/ (Sept. 2017).

131. Hossain, M. S. B. *et al. UCF-MultiOrganPath:A Benchmark Dataset of Histopathologic Images for Deep Learning-Based Organ Classification* Nov. 2024.

132. Khorfan, K. *et al. Meta-analysis to identify emerging biomarkers of pancreatic adenocarcinoma pathogenesis.* 2018.

133. Khorfan, K. *et al. Elucidating the pathogenesis of esophageal adenocarcinoma through meta-analysis of public data.* in *Journal of Clinical Oncology* **37** (American Society of Clinical Oncology, 2019), 71–71.

134. Kim, S. *et al.* Planning grant: Crowdsourcing and improving assessment of skin disease in pediatric rheumatology. *Pediatric Rheumatology* (2018).

135. Li, Y. Y. R. Y. Y. R. *et al.* Rare copy number variants in over 100,000 European ancestry subjects reveal multiple disease associations. *Nature Communications* **11,** 1–9. ISSN: 20411723. PMID: 31937769. https://pubmed.ncbi.nlm.nih.gov/31937769/ (1 Dec. 2020).

136. Lituiev, D. S. *et al.* Automatic Labeling of Special Diagnostic Mammography Views from Images and DICOM Headers. *Journal of Digital Imaging* **32,** 228–233. ISSN: 1618727X. PMID: 30465142 (2 Apr. 2019).

137. Lozano, D., Raza, L., Aljabban, J. & Hadley, D. *Meta-analysis of Gene Expression in Early-onset Alzheimers Disease (2495)* in (Wolters Kluwer Health, Inc. on behalf of the American Academy of Neurology, 2020). PMID: 32289141. https://doi.org/10.1212/wnl.94.15_supplement.2495.

138. Paik, H. *et al.* Tracing diagnosis trajectories over millions of patients reveal an unexpected risk in schizophrenia. *Scientific data* **6,** 201. ISSN: 20524463. PMID: 31615985 (1 Dec. 2019).

139. Syed, S. *et al. Identifying Potential Drug Targets for Sickle Cell Disease through Gene Expression and Pathway Analysis of GEO Data* in *Blood* **136** (Nov. 2020), 15–16.

140. Waggoner, J. J. *et al.* Improved serotype-specific dengue virus detection in Trinidad and Tobago using a multiplex, real-time RT-PCR. *Diagnostic Microbiology and Infectious Disease* **81,** 105–106. ISSN: 18790070. PMID: 25533614. https://pubmed.ncbi.nlm.nih.gov/25533614/ (2 Feb. 2015).

141. Ben-Sasson, E., Chiesa, A., Genkin, D., Tromer, E. & Virza, M. *Zerocash: Decentralized Anonymous Payments from Bitcoin* in *2014 IEEE Symposium on Security and Privacy* Original IEEE publication introducing the Zerocash protocol for privacy-preserving transactions using zk-SNARKs. (IEEE, San Jose, CA, USA, 2014), 459–474. https://doi.org/10.1109/sp.2014.36.

142. Aljabban, J. *et al.* Meta-analysis Demonstrates Altered Antigen Presentation and Maladaptive Immune Response in Chronic Urticaria. *Journal of Allergy and Clinical Immunology* **145,** AB202. ISSN: 00916749. https://doi.org/10.1016/j.jaci.2019.12.306 (2 Feb. 2020).

143. Hakonarson, H., Hadley, D., Wu, Z.-L. & Glessner, J. *Genetic Alterations Associated with Autism and the Autistic Phenotype and Methods of Use Thereof for the Diagnosis and Treatment of Autism* US Patent App. 14/131,359. 2014. https://patents.google.com/patent/US20140187447A1/en.

144. Voskamp, S. M. *et al.* Meta-analysis reveals differential gene expression in tetralogy of Fallot versus controls. *Birth defects research* **116,** e2293. ISSN: 2472-1727. PMID: 38146097 (1 Jan. 2024).

145. Wang, K. *et al.* PennCNV: An integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Research* **17,** 1665–1674. ISSN: 10889051. PMID: 17921354 (11 Nov. 2007).

146. Hossain, M. S. B. *et al. A Public Dataset of histopathology Images for deep learning model based classification* in *UCF Wellness, Education, AI, Lifestyle in Health (WEALTH) Symposium* (2022). https://stars.library.ucf.edu/wealth/2022/thursday/2/.

147. Rohr, M. W., Beardsley, J., Nakkina, S. P., Hadley, D. & Altomare, D. *Abstract 2689: FGF19 is a novel serum colorectal cancer biomarker that exerts endocrine paraneoplastic effects on hepatic tissue* in *Cancer Research* **82** (American Association for Cancer Research, June 2022), 2689–2689.

148. Hasan, L. K. *et al.* Metaanalysis Reveals Genetic Correlates of Osteoporosis Pathogenesis. *The Journal of Rheumatology* **48,** 940–945. ISSN: 0315-162X. PMID: 33262303 (6 June 2021).

149. Liu, X. *et al.* Reporting guidelines for clinical trials evaluating artificial intelligence interventions: CONSORT-AI. *BMJ* **370,** m3164 (2020).

150. Bossuyt, P. M. *et al.* STARD 2015: An Updated List of Essential Items for Reporting Diagnostic Accuracy Studies. *BMJ* **351,** h5527. https://doi.org/10.1136/bmj.h5527 (2015).

151. Pentland, A. *et al. A New Deal on Data* 2014. https://hbr.org/2014/10/a-new-deal-on-data.

152. Panahiazar, M. *et al.* Large scale advanced data analytics on skin conditions from genotype to phenotype. *Informatics* **5,** 39. ISSN: 22279709 (4 2018).

153. Von Elm, E. *et al.* The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement. *PLoS Med.* **4,** e296. https://doi.org/10.1371/journal.pmed.0040296 (2007).

154. Ogrinc, G. *et al.* SQUIRE 2.0 (Standards for QUality Improvement Reporting Excellence). *BMJ Quality & Safety* **24,** 783–791. https://pubmed.ncbi.nlm.nih.gov/26369893/ (2015).

155. Trivedi, H. M. *et al.* Large Scale Semi-Automated Labeling of Routine Free-Text Clinical Records for Deep Learning. *Journal of Digital Imaging* **32,** 30–37. ISSN: 1618727X. PMID: 30128778 (1 Aug. 2019).

156. Schulz, K. F., Altman, D. G. & Moher, D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *BMJ* **340,** c332. https://doi.org/10.1136/bmj.c332 (2010).

157. Guyatt, G. H. *et al.* GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* **336,** 924–926. https://doi.org/10.1136/bmj.39489.470347.ad (2008).

158. National Cancer Institute. *SEER Program Overview* 2025. https://seer.cancer.gov/about/overview.html.

159. National Science Foundation. *NSF I-Corps Program Overview* 2025. https://www.nsf.gov/icorps.

160. Congress, U. *Plain Writing Act of 2010* 2010. https://www.congress.gov/bill/111th-congress/house-bill/946.

161. Canales, M. T. *et al.* As-Needed Blood Pressure Medication and Adverse Outcomes in VA Hospitals. *JAMA internal medicine* **185,** 52–60. ISSN: 2168-6114. PMID: 39585709 (1 Jan. 2025).

162. Chalkia, D. *et al.* A mitochondrial bioenergetic hypothesis for autism spectrum disorder (570.3). *The FASEB Journal* **28,** 570.3. ISSN: 0892-6638 (1 Supplement 2014).

163. Chalkia, D. *et al.* Association Between Mitochondrial DNA Haplogroup Variation and Autism Spectrum Disorders. *JAMA psychiatry* **74,** 1161–1168. ISSN: 21686238 (11 Nov. 2017).

164. Hadley, D. *et al.* Analysis of six genetic risk factors highly associated with AMD in the region surrounding ARMS2 and HTRA1 on chromosome 10, region q26. *Investigative Ophthalmology and Visual Science* **51,** 2191–2196. ISSN: 01460404. PMID: 19933195 (4 Apr. 2010).

165. Kido, T. *et al.* Are minor alleles more likely to be risk alleles? *BMC Medical Genomics* **11,** 3. ISSN: 17558794. https://pubmed.ncbi.nlm.nih.gov/29351777/ (1 Dec. 2018).

166. Panahiazar, M., Baygoui, R. & Hadley, D. *A Multidimensional Gender-Based Study on UCSF Electronic Medical Record to Improve Women Health* in *Journal of Womens Health* **28** (2019), 19.

167. Lituiev, D. S. *et al.* Automated Localization and Segmentation of Mononuclear Cell Aggregates in Kidney Histological Images Using Deep Learning. *Microbiology, Immunology and Pathology* **2** (1 Feb. 2021).

168. Agency for Healthcare Research and Quality. *Health Literacy Universal Precautions Toolkit* 2020. https://www.ahrq.gov/health-literacy/improve/precautions/toolkit.html.

169. National Cancer Institute. *PRO-CTCAE™: Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events* 2025. https://healthcaredelivery.cancer.gov/pro-ctcae/.

170. Cella, D. *et al.* PROMIS: Patient-Reported Outcomes Measurement Information System. *Med. Care* **45,** S3–S11. PMID: 17443116 (2007).

171. National Library of Medicine. *About PubMed* 2025. https://pubmed.ncbi.nlm.nih.gov/about/.

172. RE-AIM.org. *RE-AIM Framework Overview* 2025. https://www.re-aim.org/.

173. Ding, Y. *et al.* A deep learning model to predict a diagnosis of Alzheimer disease by using 18 F-FDG PET of the brain. *Radiology* **290,** 456–464. ISSN: 15271315. PMID: 30398430 (3 Feb. 2019).

174. Rappoport, N. *et al.* A genome-wide association study identifies only two ancestry specific variants associated with spontaneous preterm birth. *Scientific Reports* **8,** 226. ISSN: 20452322. PMID: 29317701 (1 Dec. 2018).

175. Aljabban, J. *et al. Mo1125 GENETIC ANALYSIS OF ULCERATIVE COLITIS RESPONSE TO INFLIXIMAB* in *Gastroenterology* **158** (WB Saunders, May 2020), S-796-S–797. https://doi.org/10.1016/s0016-5085(20)32666-4.

176. Allen, D. Z. *et al.* Meta-Analysis illustrates possible role of lipopolysaccharide (LPS)-induced tissue injury in nasopharyngeal carcinoma (NPC) pathogenesis. *PloS one* **16** (ed Mummidi, S.) e0258187. ISSN: 1932-6203. PMID: 34648530 (10 Oct. 2021).

177. Archive, T. C. I. *TCIA: Images for Cancer Research* 2025. https://www.cancerimagingarchive.net/.

178. National Center for Biotechnology Information. *ClinVar* 2025. https://www.ncbi.nlm.nih.gov/clinvar/.

179. National Center for Biotechnology Information. *dbGaP: Database of Genotypes and Phenotypes* 2025. https://www.ncbi.nlm.nih.gov/gap.

180. Broad Institute. *gnomAD: Genome Aggregation Database* 2025. https://gnomad.broadinstitute.org/.

181. Snyder, C. F. *et al.* Implementing Patient-Reported Outcomes in Routine Oncology Care. *J. Clin. Oncol.* **30,** 4299–4304. https://pubmed.ncbi.nlm.nih.gov/22048932/ (2012).

182. Hadley, D., Pan, J., Sirota, M., Chen, B. & Oskotsky, B. Search Tag Analyze Resource ( STAR ): An online platform to crowd-source genomic disease signatures from open digital samples . *AMIA Joint Summits on Translational Science proceedings AMIA Summit on Translational Science* **In press,** 1–5. PMID: 26306233 (2015).

183. American Medical Association. *Current Procedural Terminology (CPT)* 2025. https://www.ama-assn.org/practice-management/cpt.

184. Hadley, D. *Crowdsourcing an Open COVID-19 Chest Radiograph Imaging Repository for Artificial Intelligence Research* 2022. nct: NCT05384912.

185. Aljabban, J. *et al. Assessing the potential of immunotherapy in treating chronic lymphocytic leukemia through meta-analysis.* in *Journal of Clinical Oncology* **37** (American Society of Clinical Oncology, 2019), 7531–7531.

186. Aljabban, J. *et al. Characterization of Monoclonal Gammopathy of Undetermined Significance Progression to Multiple Myeloma through Meta-Analysis of GEO Data* in *Blood* **134** (American Society of

Hematology Washington, DC, Nov. 2019), 4395–4395.

187. Arora, A. *et al.* Predictive Models for Length of Stay and Discharge Disposition in Elective Spine Surgery: Development, Validation, and Comparison to the ACS NSQIP Risk Calculator. *Spine.* https://www.semanticscholar.org/paper/8773e3620f6ed9ef3a0d6781c7e713a305c8038f (2022).

188. Lituiev, D., Cha, S. J., Sohn, J. H., Hadley, D. & Laszik, G. Z. *A Deep Learning Model Can Identify Rejection in Transplant Kidney Biopsies* in *Laboratory Investigation* **100** (2020), 1472.

189. Monticciolo, D. L. *et al.* Breast Cancer Screening in Women at Higher-Than-Average Risk. *J. Am. Coll. Radiol.* **15,** 408–414 (2018).

190. Orlin, A. *et al.* Association between high-risk disease loci and response to anti-vascular endothelial growth factor treatment for wet age-related macular degeneration. *Retina* **32,** 4–9. ISSN: 0275004X. PMID: 21878851. https://pubmed.ncbi.nlm.nih.gov/21878851/ (1 Jan. 2012).

191. Panahiazar, M. & Hadley, D. *Large-Scale Analysis of UCSF Electronic Medical Record to Improve Women's Health in Breast Cancer and Cardiovascular Diseases* in *Journal of Womens Health* **27** (2018), 14.

192. Panahiazar, M. *et al. A Trend of Eight-Years Big Data Analytics of Electronic Medical Records to Review and Study Diagnosis and Treatment of Coronary Artery Disease in Different Genders* Apr. 2021.

193. Panahiazar, M. *et al.* Gender-based time discrepancy in diagnosis of coronary artery disease based on data analytics of electronic medical records. *Frontiers in cardiovascular medicine* **9,** 969325. ISSN: 2297-055X. PMID: 36505372 (2022).

194. Sanna-Cherchi, S. *et al. a High Frequency of Genomic Disorders in Patients With Congenital Kidney Malformations* in *Pediatric Nephrology* **27** (2012), 1621–1622.

195. Sohn, J. H. *et al.* Prediction of Future Healthcare Expenses of Patients from Chest Radiographs Using Deep Learning: A Pilot Study. *Scientific reports* **12,** 8344. ISSN: 2045-2322. PMID: 35585177 (1 May 2022).

196. Hadley, D. *et al.* in *Leveraging Technology as a Response to the COVID Pandemic* (ed Pappas, P. H. F. H. P.) 1st ed., 151–180 (Productivity Press, Nov. 2022). ISBN: 9781003352297.

197. Hadley, D. *et al.* The impact of the metabotropic glutamate receptor and other gene family interaction networks on autism. *Nature Communications* **5,** 4074. ISSN: 20411723. PMID: 24927284 (1 Sept. 2014).

198. SMART Health IT. *SMART App Launch Implementation Guide* HL7 FHIR Implementation Guide describing OAuth2/OIDC app launch for FHIR apps. 2023. https://www.hl7.org/fhir/smart-app-launch/.

199. Hadley, D. *et al.* Patterns of sequence conservation in presynaptic neural genes. *Genome Biology* **7,** R105. ISSN: 14747596. PMID: 17096848 (11 Jan. 2006).

200. U.S. Food and Drug Administration. *Unique Device Identification (UDI) System* 2024. https://www.fda.gov/medical-devices/unique-device-identification-system-udi-system.

201. Sanna-Cherchi, S. *et al.* Copy-number disorders are a common cause of congenital kidney malformations. *American Journal of Human Genetics* **91,** 987–997. ISSN: 00029297. PMID: 23159250 (6 Dec. 2012).

202. Aljabban, J. *et al.* Characterization of monoclonal gammopathy of undetermined significance progression to multiple myeloma through meta-analysis of GEO data. *Heliyon* **9,** e17298. ISSN: 2405-8440. PMID: 37539132 (7 July 2023).

203. Bucan, M. *et al.* Genome-wide analyses of exonic copy number variants in a family-based study point to novel autism susceptibility genes. *PLoS Genetics* **5** (ed

Gibson, G.) e1000536. ISSN: 15537390. PMID: 19557195 (6 June 2009).

204. American Statistical Association. ASA Statement on Differential Privacy. *Amstat News.* https://www.amstat.org/asa/files/pdfs/POL-DifferentialPrivacyStatement.pdf (2021).

205. Kay, D., Farhangi, A., Guo, Z., Castiglioni, A. & Hadley, D. *Can Artificial Intelligence Address The Burden Associated With Scoring Narrative Assessments?* in *Medical Science Educator* (Springer, Jan. 2023), 1–35.

206. Sleiman, P. *et al.* GWAS meta analysis identifies TSNARE1 as a novel Schizophrenia / Bipolar susceptibility locus. *Scientific Reports* **3,** 3075. ISSN: 20452322. PMID: 24166486 (1 Jan. 2013).