📄 PRD: ChainIndexed – On-Chain H-Index with AI Citation Verification

🚨 Problem Statement & Motivation

💸 The Gatekeeping of Scientific Progress

The current scientific publishing ecosystem is heavily centralized and profit-driven, dominated by corporations like **Elsevier, Springer Nature, and Wiley**, which:

- Charge exorbitant subscription fees to access knowledge—locking out the global south and independent researchers.

- Demand publishing fees (often $3,000–10,000 per paper) for open access—even when research is publicly funded.

- Enforce **closed citation systems**, limiting visibility and discoverability of research across domains.

🙍‍♀️ Misattribution and Credit Hogging

- First authors or lead institutions often overshadow the work of collaborators or junior researchers.

- Citation gaming, "blank citations" (fake references), and collusion distort genuine scientific merit.

- H-index and similar metrics are opaque, privately computed, and sometimes manipulated.

🎯 Product Overview

ChainIndexed is a decentralized protocol that enables researchers to register their scientific contributions on-chain and automatically compute a **verifiable H-index**, using **Chainlink**

Functions and **AI-based citation validation**. The system ensures that only legitimate, referenced scientific work contributes to a user's academic impact score.

🧠 ChainIndexed: What This Fixes

✅ Decentralized & Transparent Impact Measurement

- Researchers control their contributions: **no paywalls, no gatekeepers**.

- H-index and citation counts are **verifiable**, **AI-validated**, and **publicly auditable** on-chain.

✅ Fair Attribution & Anti-Gaming

- Only **verified citations** contribute to scores—ensured via AI context checks.

- All submissions and updates are timestamped, immutable, and tied to provable wallet identities (or ORCID).

✅ Global Accessibility

- Anyone with a wallet can register, track, and showcase research—enabling participation across geographies and disciplines.

🧩 MVP Goal

Build an MVP that allows:

- A user (researcher) to submit their paper and the metadata of their research contribution.

- Chainlink Functions to summarise and index the content of the paper using external APIs and an LLM.

- A smart contract to create/update a verified H-index on-chain.

🛠️ Tech Stack

Layer	Tools / Technologies
Smart Contract	Solidity, Hardhat (or Foundry)
Oracle Layer	**Chainlink Functions**, Chainlink Automation (optional)
Off-Chain AI/Logic	OpenAI (or any public LLM), Citation APIs (Crossref, Semantic Scholar)
Storage	IPFS (for submission metadata, optional)
Chain	Polygon Mumbai or Avalanche Fuji (for low gas and Chainlink compatibility)
Hosting	IPFS

🔑 Core Features (MVP)

1. 📚 Contribution Submission

- Researchers can submit:

    - Title

    - Abstract

    - CID or DOI

    - Optional author ID or ORCID

2. 🤖 AI-Verified Citation Checker

- Chainlink Function fetches citation data from Crossref/Semantic Scholar.

- Uses OpenAI (or similar) to validate that the citing article meaningfully refers to the original work.

- Returns a filtered citation count.

3. 📈 H-Index Calculator (on-chain)

- Smart contract calculates H-index based on verified citation counts.

- Stores updated value in a public mapping: `mapping(address => uint256) public hIndex`.

5. 🔐 Basic Anti-Spam Controls

- Users must stake a small amount of test tokens to submit research (prevents spam).

- Only submissions with valid format (CID or DOI) are processed.

🔮 Future Releases (Post-MVP)

🔁 Scheduled Citation Re-checks

- Use Chainlink Automation to re-fetch citation data weekly and auto-update scores.

🥇 Tokenized Reputation Layer

- Introduce `SCI` token to reward contributors and reviewers.

- Badges/NFTs based on citation thresholds.

📖 Research Review Module

- Allow peers to review/endorse research on-chain.

- Weight H-index based on community trust scores.

📊 Citation Graph + Provenance

- Visualize citation relationships between on-chain contributions.

- Use a DAG model or IPFS-linked citation trees.

🌐 Decentralized Search Oracle

- Use Chainlink Functions to index submissions and provide a search API for research topics or keywords.

🧠 AI-Enhanced Metrics

- Detect novelty of research (via embeddings + cosine similarity).

- Summarize contributions with LLMs on-chain or via Functions.

📋 Key Smart Contract Interfaces

solidity
CopyEdit
```solidity
function submitResearch(string calldata doiOrCid, string calldata
abstract, string calldata title) external;
function updateHIndex(address researcher, uint256 newHIndex) external
onlyFunctionNode;
function getResearcherData(address researcher) public view returns
(Research[] memory, uint256 hIndex);
```

📅 Development Timeline (for Hackathon)

Day	Task
1	Set up contract + research submission logic
1	Write Chainlink Function: citation API + LLM scoring
1.5	Test end-to-end H-index update flow
2	Build frontend dashboard + record demo

✅ Success Metrics

- At least 2–3 valid researcher profiles with working H-index scores.

- Accurate citations verified via AI (log confidence scores).

- 1-click demo flow from submission → H-index → dashboard.

🌍 Why This Matters (Beyond the Hackathon)

- Builds infrastructure for **DeSci (Decentralized Science)**—a growing global movement to democratize scientific knowledge.

- Empowers **researchers over institutions**, incentivizing open collaboration and credit fairness.

- Could plug into reputation systems, grant eligibility platforms, or DeSci DAOs in the future.