

CURATOR RESEARCH PRIVATE LIMITED

Evidence-bound AI and knowledge systems, built in India

For regulatory and document-heavy work: software, models and connected knowledge, with privacy, cost, evaluation and human authority built in.

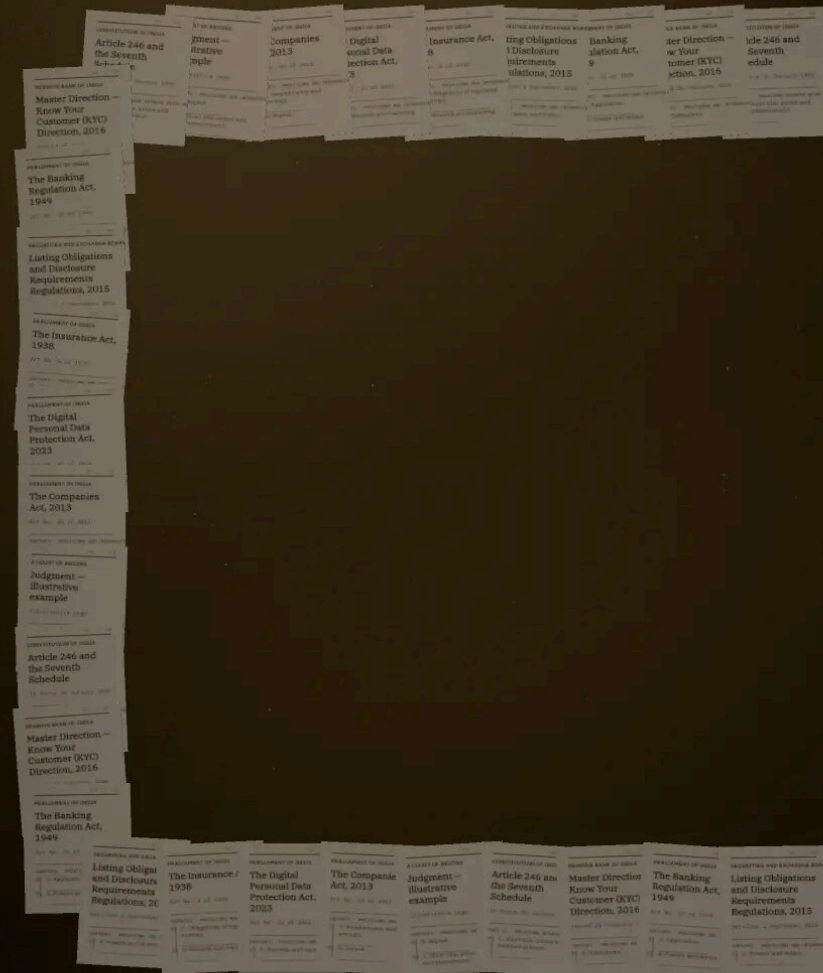
Incorporated 24 August 2026 · Panchkula, Haryana

CIN U62011HR2026PTC149626 · Udyam UDYAM-HR-13-0040757

www.curatorresearch.com · founder@curatorresearch.com

Supporting document · DPIIT StartUp Recognition · October 2026

The declared edge — where the evidence stops

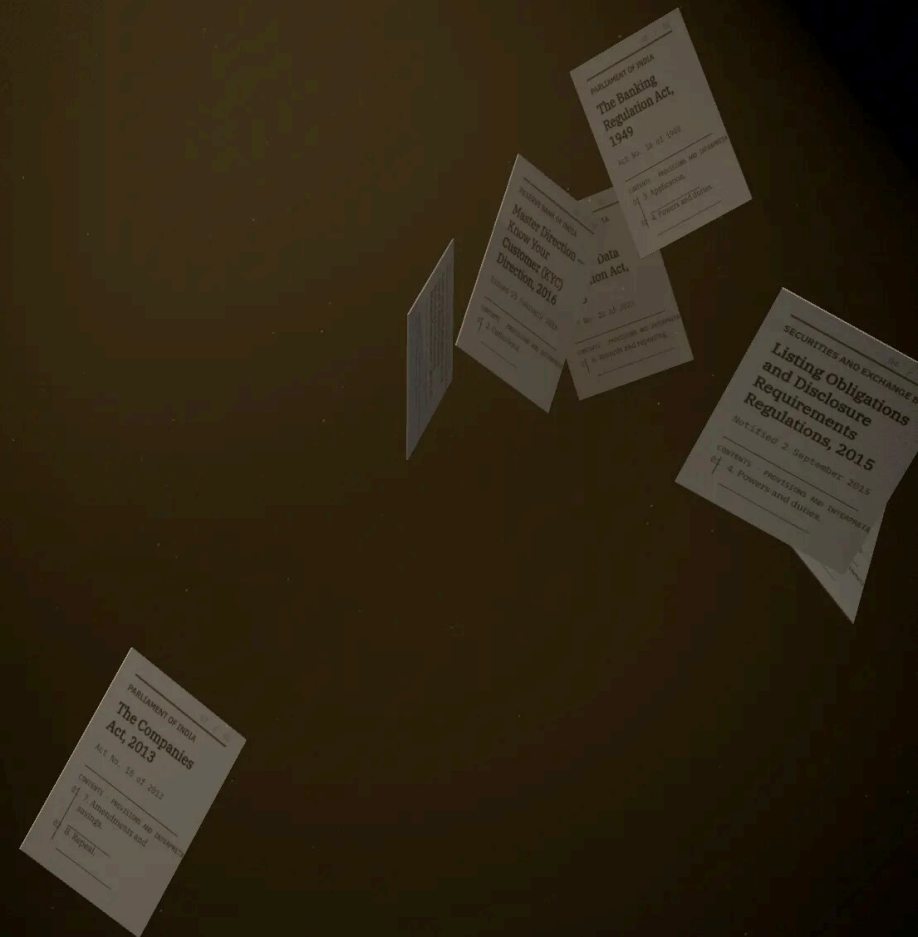


THE PROBLEM

Regulatory change crosses documents and systems

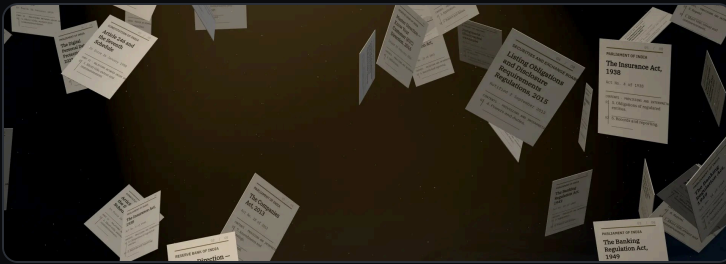
- **Regulation is being restructured.** Regulators are consolidating and rewriting large bodies of instructions; a policy that cites an older instrument can point somewhere that has moved.
- **The work is manual and costly.** Teams read PDFs and keep spreadsheets to say which rule applied on a date, what changed and what it affects, while compliance costs keep rising.
- **Provenance is required.** Every regulatory conclusion must trace to its source, version and effective date.
- **Fluent is not accountable.** An answer that cannot be traced to its source cannot be relied on before a regulator, an auditor or a court.

Acts · directions · circulars · judgments · filings



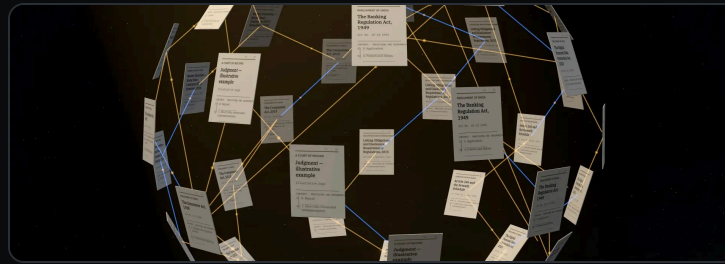
OUR APPROACH

Connected sources and accountable workflows



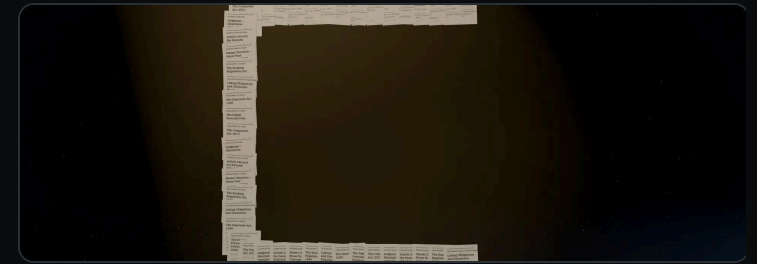
Scattered information

Circulars, directions, Acts, judgments and filings: many sources, many formats, each changing on its own timetable.



Connected knowledge

We connect each source to what it amends, replaces, cites and depends on, with the date each link held.



One accountable system

Our systems show their sources and limits, state when evidence is insufficient, and keep consequential actions under human authority.

FROM DOCUMENTS TO CITED ANSWERS

01

Sources

Acts, regulations, directions, circulars, orders and judgments



02

Collect and parse

A self-scheduling fleet; tables, annexures, scans and bilingual pages, with lineage



03

Enrich and link

Entities, facts and citations; links that amend, replace, cite and interpret, each dated



04

Retrieve and rerank

Semantic, relational and graph-traversal search, reranked by our trained models



05

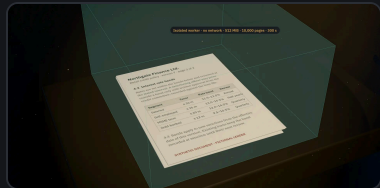
Cited synthesis

Answers tied to the exact clause; the system abstains when support is insufficient

Receipts and records of what was done, what was searched and what actually ran, at every step

THE PRODUCT FAMILY

Five products built around shared evidence



Patra

Documents you can prove: PDF tools with a receipt of what was done to every file.

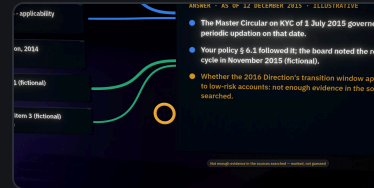
● Built · internal testing



BanyanGraph

A deeply connected knowledge base of Indian law and regulation.

● In internal use



Curator

Deep research you can defend: cited, date-correct answers.

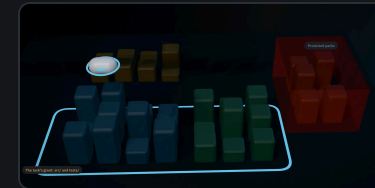
● Interface in development



ChangeProof

When the law changes, everything it touches.

● MVP · internal evaluation



Karta

An AI coding harness with local-first agents.

● Internal evaluation

Together, the five products are designed to take document-heavy industries and professional firms, from BFSI to legal, CA and CS practices, from source to decision: documents with a record of what was done, a dated knowledge base of Indian law and regulation, cited answers for the date asked, regulatory change traced to everything it touches, and governed implementation.



BANYANGRAPH

A deeply connected knowledge base of Indian law and regulation

- **A chain of authority.** The Constitution, the Acts it empowers, regulators' instruments and the judgments that interpret them.
- **Dated links.** Instruments connect by what they empower, amend, replace, cite and interpret. Each link keeps when it held and when we recorded it.
- **Parts you can follow.** Chapters, clauses, tables and defined terms, searched by meaning and followed through the graph.

Collection framework: collectors for 16 Indian regulators and authorities (RBI, SEBI, IRDAI, PFRDA, IBBI, IFSCA and more), 13 courts and tribunals, Parliament and the Gazette of India.

● In internal use (snapshot 27 Aug 2026) · customer search in development

Documents you can prove

- **Inspected first.** Pages, encryption, tables and risk flags are checked before anything is done.
- **Isolated by design.** Each operation runs in its own worker: no network, bounded memory, pages and time.
- **Real redaction.** Regions a person marked are burned into the page, not covered over.
- **One receipt.** Every step, its digests, and whether anything was kept.

Convert, compress, redact, merge, split and re-order, OCR in English and Hindi, protect and unlock, watermark and compare.

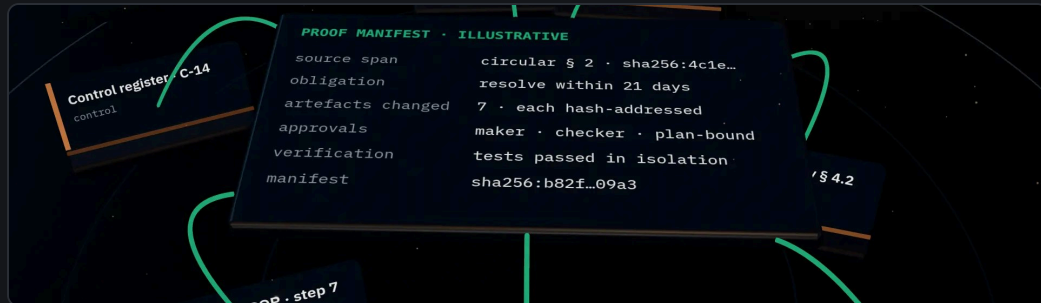
● 31 operations built and tested (25 Sep 2026) · hosted public version in development

A REAL RECEIPT · SYNTHETIC DOCUMENT

Operation	markdown.to_pdf
Operation version	1.0.0
Determinism	D1 Same input, same bytes out.
Fidelity	reconstructed Rebuilt, not preserved. Patra says so itself.
Ran on	isolated-worker
Source digest	b49ec949...3601c04
Output digest	7f4efa28...2096b58
Duration	234 ms
Retained	no The document was not kept.

Patra CLI 0.1.0, 11 Sep 2026: `patra markdown.to_pdf sample.md -o sample.pdf --receipt receipt.json` on a synthetic two-clause board resolution. Values copied unchanged; published on www.curatorresearch.com/products/patra.

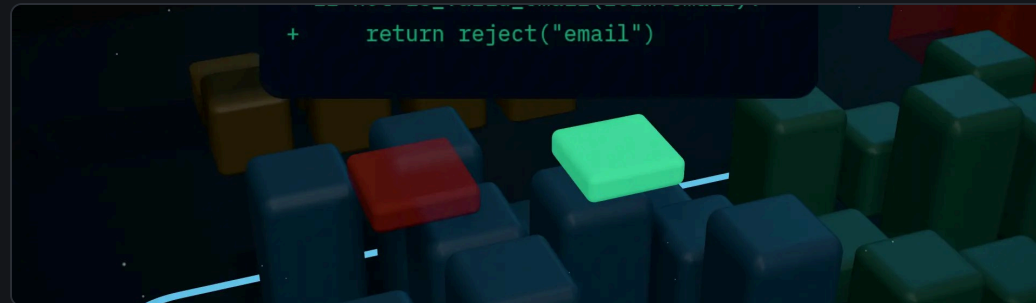
Regulatory change and governed implementation



ChangeProof

ChangeProof is designed to trace regulatory requirements to affected policies, procedures, controls and code. It connects functional obligations to technical implementation, with named owners, approval of the exact plan and a record of the work performed.

- A specified MVP under internal evaluation



Karta

An AI coding harness with local-first agents: local models first, inside the permissions a task allows, on a cost-governed route, stopping for a person where it should, with a record of every step. Protected paths stay fenced until a person approves, and commands run in a sandbox with the network off by default.

- In internal evaluation; not released

Technical differentiation

Our differentiation combines dated knowledge, domain retrieval and controls for professional decisions.

1. **Deeply connected.** A proprietary ingestion pipeline extracts deep-linked knowledge across statutes, directions, documents and facts, with provenance, lineage and both when a rule held and when we recorded it.
2. **Retrieval built for regulation.** Semantic, relational and graph-traversal search, reranked by our trained models, keeps irrelevant hits and unsupported answers low; the system abstains when support is insufficient.
3. **Domain models we host and publish.** Fine-tuned enrichment, reranking, synthesis and orchestration models run on our own infrastructure; 14 are public experimental releases on Hugging Face (version 0.1.0).
4. **Accountable.** Receipts record what actually ran; consequential changes are approved by named owners against the exact plan shown.

PUBLISHED WORK

Models we have built and published

14 public experimental releases · version 0.1.0 · huggingface.co/curator-research

Structured enrichment (8B, thinking on)

On 48 held-out documents: valid JSON 48/48, schema 0.970, grounding 0.877 and semantic F1 0.559.

Reranker

On 432 held-out query groups (112 distinct queries), nDCG@10 0.945 versus 0.918 for its base (95% CI +0.016 to +0.040) and MRR@10 0.959 versus 0.904.

Workflow orchestration (8B, thinking on)

Correct next action on 31/48 held-out turns, against 22/48 for the best fixed action.

Grounded synthesis (thinking mode)

On 118 questions with cited references, raw citation recall was 0.918 versus the teacher's 0.875; the teacher led on raw gate and judge scores.

Curator's own held-out evaluations of experimental releases; not independently validated. Llama-lineage releases: Built with Llama. [Model cards and evaluation details](#)

Models	14	Filter by name	Base only	Inference Available	Sort: Trending
 curator-research/Llama-3.3-Qwen3-8B-Curator-0.1.0-Enrichment-Thinking-ON	Updated 17 minutes ago	↓ 10			
 curator-research/Llama-3.3-MiniLM-L6-v2-Curator-0.1.0-Reranker-Multisource	Text Classification · 22.7M · Updated about 12 hours ago	↓ 2			
 curator-research/Llama-3.3-Qwen3-8B-Curator-0.1.0-Enrichment-Thinking-OFF	Updated 17 minutes ago	↓ 9			
 curator-research/Qwen3-8B-Curator-0.1.0-Orchestration-Thinking-ON	Updated 16 minutes ago	↓ 10			
 curator-research/Qwen3-8B-Curator-0.1.0-Orchestration-Thinking-OFF	Updated 16 minutes ago	↓ 10			
 curator-research/Llama-3.3-Qwen3-8B-Curator-0.1.0-Synthesis-OFF	Updated 16 minutes ago	↓ 14			
 curator-research/Llama-3.3-Qwen3-8B-Curator-0.1.0-Synthesis-THINK	Updated 16 minutes ago	↓ 14			
 curator-research/Llama-3.3-Qwen3-8B-Curator-0.1.0-Synthesis-AGENTIC	Updated 17 minutes ago	↓ 10			
 curator-research/Llama-3.3-Qwen3.6-35B-A3B-Curator-0.1.0-Enrichment-Thinking-ON	Updated 16 minutes ago	↓ 13			
 curator-research/Llama-3.3-Qwen3.6-35B-A3B-Curator-0.1.0-Enrichment-Thinking-OFF	Updated 16 minutes ago	↓ 13			
 curator-research/Qwen3.6-35B-A3B-Curator-0.1.0-Orchestration-Thinking-ON	Updated 16 minutes ago	↓ 12			
 curator-research/Qwen3.6-35B-A3B-Curator-0.1.0-Orchestration-Thinking-OFF	Updated 16 minutes ago	↓ 13			
 curator-research/Llama-3.3-Nemotron-Cascade-2-30B-A3B-Curator-0.1.0-Enrichment-Thinki...	Text Generation · 32B · Updated 17 minutes ago	↓ 5			
 curator-research/Llama-3.3-Nemotron-Cascade-2-30B-A3B-Curator-0.1.0-Enrichment-Thinki...	Text Generation · 32B · Updated 17 minutes ago	↓ 2			

HOW WE BUILT IT

From the founders' research to the company

BEFORE INCORPORATION · THE FOUNDERS' RESEARCH

From October 2025 the founders researched regulatory knowledge systems: market-data infrastructure, source collection from regulators and tribunals, a Version 1 ingestion and retrieval pipeline with graph-RAG research, and model fine-tuning and evaluation on a 60-day NVIDIA DGX Cloud Innovation Lab allocation (eight 80 GB H100 GPUs, **77.0%** mean utilisation across measured multi-day runs).

THE COMPANY

24 Aug 2026	Curator Research Private Limited incorporated in Panchkula, Haryana
Sep 2026	Patra, BanyanGraph, Curator, ChangeProof and Karta, each in development
3 Oct 2026	14 models published on Hugging Face as public experimental releases (version 0.1.0)

Version 1 demonstrated that parsing, enrichment, retrieval and specialised models could work together. Version 2 is our work to make those foundations dependable at different scales.

Customers and planned revenue

Software

For professionals and firms: Patra's document tools and Curator's research interface.

Subscriptions.

For individuals and professionals; legal, CA and CS firms, MSMEs and banks

Knowledge access

For RegTech, GRC and advisory firms, and regulated institutions' internal systems: planned licensed access to BanyanGraph knowledge.

Licensed access.

For firms and regulated organisations

Implementation services

Scoped engagements: knowledge bases, regulatory research and change workflows, model evaluation and integration inside the client's boundary.

Scoped engagements.

For organisations building a knowledge base of what they already know

First market: BFSI, a document-heavy and heavily regulated industry. Pre-revenue and self-funded; pricing is in discovery with prospective design partners, to suit individual professionals, firms and institutions.

WHY IT SCALES

Reusable knowledge and phased expansion

- **One evidence layer, many products.** The same dated, cited knowledge serves documents, research, change and governed action; source improvements can support multiple products while client records stay separated.
- **Deployable inside the client's boundary.** Verifiable no-egress document processing today; local and private knowledge operation is a direction under development.
- **A growing team.** We plan to expand in phases, bringing in technical and functional specialists as products and engagements mature.



THE COMPANY

Founded and directed by Bhavika Mukherjee and Aditi Chadha

- An Indian technology company with an all-women board.
- Self-funded; no outside investment.
- AI-native, under human authority: every task has an owner, a limit, a reviewer and a record.
- Evidence first: every figure we publish carries its date and scope.

Registrations

Curator Research Private Limited
CIN U62011HR2026PTC149626
Incorporated 24 August 2026 · Panchkula, Haryana
MSME (Udyam) UDYAM-HR-13-0040757 · Micro

Contact

www.curatorresearch.com
founder@curatorresearch.com

Illustrations in this deck are stills from www.curatorresearch.com, and on slide 9 a screenshot of the public Hugging Face model list taken on 4 October 2026; demonstration content shown in the stills is fictional. Product stages are quoted as the website publishes them.