IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Artificial Intelligence (Ai) In Restructuring Financial Reports And Their Applications In India

[1]Sahil Bhandari, [2]Shriram Ghadge, [3]Omkar Kakade, [4]Mohak Khullar, [5]Onkar Mawal

Students of Master of Business Administration (MBA), Narsee Monjee Institute of Management Studies (NMIMS), Mumbai

Abstract: Artificial Intelligence (AI)—spanning OCR, knowledge graphs, and large language models (LLMs)—is now capable of turning India's sprawling, multi-format financial data into structured, decision-grade reports at scale. This paper synthesizes the background and need, global and Indian applications, and near-term use cases across corporate finance, banks/NBFCs, taxation, investment research, insurance, and regulation. We quantify time, cost, and carbon/ESG benefits using official datasets wherever possible and provide transparent assumptions for guesstimates where official numbers don't exist. We incorporate RBI's emerging FREE-AI governance, SEBI's 2025 AI/ML consultation, GST e-invoicing scale, UPI/payments indicators, and Account Aggregator (AA) adoption to ground a pragmatic roadmap for Indian institutions. We close with implementation risks, research gaps, and a five-year outlook (FY26–FY30) for measurable national savings.

I. Context and the need for AI-driven financial data restructuring

India's financial information flows are rich but fragmented: invoices (PDFs, images), bank statements (CSV, PDF, AA APIs), GST returns, MCA filings, bureau and alternate data, and unstructured disclosures. Lenders, CFO teams, and regulators spend tens of millions of person-hours each year extracting, cleaning, reconciling, and reformatting these datasets into working capital packs, regulatory returns, and board-ready reports. The sheer scale is visible in public infrastructure metrics:

- UPI processed 14.26 billion transactions in July 2025 (and comparable scale in preceding months), with RBI's Payment System Indicators also reporting 82 million UPI QR codes and 2.21 lakh ATMs—an indication of the digital rails over which structured data can flow.
- GST e-invoicing has reached industrial scale: ~110 crore e-invoices in May 2024 (1.1 billion)—a monthly snapshot that demonstrates the volume AI can touch if embedded in the e-invoice lifecycle.
- Account Aggregator (AA) adoption has accelerated: ~₹1.6 lakh crore of loans in FY25 across 1.89 crore loan accounts were facilitated using AA-based data sharing; Sahamati also reports hundreds of institutions live and 269+ million consents fulfilled cumulatively.

These volumes translate to massive manual effort if handled item-by-item. AI can ingest messy inputs, detect entities, normalize to Indian taxonomies (GST, Ind-AS, XBRL tags), reconcile across ledgers, and draft narrative disclosures—with human-in-the-loop assurance.

From a policy standpoint, India's regulators are actively laying the governance and compute foundation:

- RBI released the committee report on FREE-AI (Framework for Responsible and Ethical Enablement of AI) on 13 Aug 2025, outlining "Seven Sutras" (principles) and multi-pillar recommendations for trustworthy AI in finance.
- SEBI published (20 Jun 2025) a Consultation Paper proposing guidelines for responsible AI/ML usage by market Infrastructure institutions and intermediaries.

- The Union Cabinet approved the IndiaAI Mission with ₹10,300+ crore and a public AI compute target of 10,000+ GPUs to democratize access to training/inference and to build domain-specific foundational models.

Bottom line: AI is needed to keep up with the scale and complexity of India's financial data. Its adoption must be governed (RBI/SEBI) and energized (IndiaAI compute), while staying grounded in local reporting regimes (GST, MCA, SEBI, RBI supervisory returns).

II. What is "financial data restructuring"?

We use the term to mean the AI-assisted transformation of raw financial artifacts into validated, interoperable data structures and decision-ready outputs:

1. Acquisition: OCR/vision models read PDFs/images (invoices, bank statements, stamped loan docs) and fetch machine data via AA or APIs.
2. Normalization: Schema mapping to Indian standards—GST sections (B2B/B2C), HSN/SAC codes, Ind-AS chart of accounts, SEBI disclosure tags, RBI return templates.
3. Reconciliation: Cross-checks among GST e-invoice, e-way bill, GSTR-1/3B, ledger entries, bank credits/debits, bureau and GSTN AA data.
4. Reasoning & Drafting: LLMs generate schedules, notes to accounts, variances, CAMs/CIMs, board MIS, and audit trails; detect anomalies.
5. Assurance: Validation rules, human-review workflows, model-risk documentation aligned to RBI FREE-AI and SEBI AI guardrails.

III. Global landscape (selected cases/tools)

- Invoice & spend AI (Coupa, Tipalti, SAP Concur, Klippa, Rossum, UiPath): automated invoice capture, 2/3-way match, GL coding, and anomaly checks.
- OCR + graph + LLM stacks: layout-aware OCR (e.g., Tesseract variants, PaddleOCR), vector DBs, and domain LLMs fine-tuned on financial forms to output XBRL/JSON.
- Banking: AI-assisted IFRS/CECL provisioning, model-risk documentation, and continual KYC/AML monitoring.
- Reg-reporting: Rule engines mapping core-banking data to Basel, liquidity returns, and EBA/EIOPA templates in other markets.

The relevance for India: these are proven workflows we can localize to GST/Ind-AS/SEBI/RBI norms with Indian language support and AA/GSTN/UPI data pipes.

IV. India landscape: what's live (2024–25)

4.1 Invoice Public rails & regulators

- UPI & digital payments (RBI Payment System Indicators, July 2025) show persistent hyper-scale—critical for low-friction data acquisition (e.g., cash-flow analytics, fraud models).
- GST e-invoice volumes—~110 crore e-invoices in May 2024—support end-to-end AI flows: extraction → vendor master checks → GST reconciliation → accruals.
- Account Aggregator: system-wide scale in FY25 (₹1.6 lakh crore loans, 1.89 crore loan accounts; hundreds of institutions live).
- RBI FREE-AI (Aug 2025) and SEBI AI Consultation (Jun 2025) set the governance expectations (human oversight, audits, fairness, explainability).
- IndiaAI Mission brings public compute and domain LMMs—vital for multilingual OCR/NLP on financial artifacts across Bharat.

4.2 Industry examples

- Perfios CAM AI: AI-generated credit assessment memos; press coverage cites ~85% faster automation, 40% error reduction.
- Clear (ClearTax): Microsoft case study shows Azure OpenAI-powered tax assistants/chat on WhatsApp at 200,000+ users, illustrating LLM adoption in compliance workflows.
- Tally's "zero-entry" vision: AI-driven document → ledger conversion and auto-categorization to shrink bookkeeping effort.

These signals show Indian BFSI and compliance ecosystems are already piloting AI restructuring—often on top of DPI rails (AA, UPI, GSTN).

V. DEEP-DIVE USE CASES (NEAR-TERM, INDIA)

5.1 Corporate finance / controllership

- AP/AR automation: OCR + entity matching; 2-/3-way matching; GST input tax credit checks; e-invoice–ledger reconciliation; automated accruals.
- Period close acceleration: ML-based variance analysis; LLM summaries for audit committee packs; auto-drafting notes to accounts.
- Treasury cash visibility: AA-based bank data ingestion; cash forecasting; hedging workflows with anomaly and counterparty risk alerts.

5.2 Banking & NBFCs

- CAM automation (lender packs) and working-capital assessment on AA bank data + GST flows.
- Early warning systems: payment irregularities, GST filings slippage, bounce/overdue patterns; tie-ins with RBI's supervisory focus on financial stability.
- Collections analytics & restructuring suggestions grounded in borrower cash-flows and seasonality.

5.3 Taxation (GST & Direct Tax)

- E-invoice ingestion → GSTR-1/3B prep → ITC validation; AI flags mismatches and suspicious vendor chains (supports CBIC/PIB' s anti-fake ITC drive).
- Income-tax prefill & anomaly checks (building on CBDT's faceless and prefill initiatives); TIN 2.0, updated returns scale highlighted in Year Ender 2024.

5.4 Investment research & securities markets

- LLM co-pilots to parse offer documents, call transcripts, and filings; cross-tag to SEBI disclosure requirements; bias /misinformation safeguards per SEBI's 2025 consultation.

5.5 Insurance

- AI reading of proposals/medical reports, fraud flags; automated regulatory reporting to IRDAI templates.

5.5 Regulators & market infrastructure

- Suptech (regulatory AI): outlier detection from filings and transactions; explainable models with audit trails aligned to FREE-AI.

VI. QUANTIFYING THE IMPACT FOR INDIA: TIME, COST, CARBON/ESG

We use a mix of (a) official statistics to anchor scale and (b) defensible assumptions to model benefits. Where ranges are wide, we show conservative values.

6.1 PROCESS-LEVEL TIME SAVINGS (PER 1 MILLION ITEMS)

We model typical Indian finance processes and compute FTE savings per 1 million transactions based on before/after minutes.

6.1.1 Results (per 1M items):

- Invoice capture (ap): ~28–29 FTE-years saved.
- Bank reconciliation: ~18 FTE-years saved.
- GST return prep & match: ~36–37 FTE-years saved.
- Expense audit: ~19–20 FTE-years saved.
- AR cash application: ~27–28 FTE-years saved.

These are order-of-magnitude guides for business cases; actuals vary by data quality and ERP maturity.

6.1.2 National e-invoicing scenario, FY26–FY30 (conservative)

Anchor: official 110 crore e-invoices in may 2024 (~1.1b monthly snapshot). We conservatively model 1.5b → 2.2b e-invoices/year in fy26→fy30 with ai embedded at ingestion and matching.

6.1.3 Assumptions (transparent):

- Time saved = 4 minutes/invoice (capture + validations + posting)
- Cost saved = ₹10/invoice (blended labor/paper/rework)
- Work year = 2,000 hours (8h×250 days)

6.1.4 Outputs (model):

- Hours saved: grows from ~100 million hours (FY26) to 147 million hours (FY30).
- FTE saved: rises from ~50,000 FTE-years to ~74,000 FTE-years.
- Rupee savings: ~₹15,000 crore (fy26) to ₹22,000 crore (FY30) (conservative).

Why conservative? Many enterprises today still spend more than 4 minutes per invoice end-to-end, and error-induced rework can be substantial. At the same time, some large shared-service centers are already faster; hence we show central estimates and not aggressive best-case numbers.

6.2 CARBON/ESG IMPACT

We combine (i) official activity indicators and (ii) published studies on digital payments and ATM energy to approximate carbon effects.

6.2.1 Digital payments substitution effects:

A 2024 PwC India paper on digital payments' climate impact estimates that (a) slower PoS expansion and (b) shifts to UPI/QR together yielded ~30,000 tCO_2 savings through FY21–FY23 (≈10k + 20k). We summarize those and add a modeled incremental FY26–FY30 potential as UPI scales further.

6.2.2 ATM energy use

Peer-reviewed/technical literature puts typical Indian ATM electricity use at ~2.4 MWh/year (excluding air-conditioning, depending on climate). Fewer cash withdrawals and cash-lite journeys reduce the need for physical trips and the supporting infra over time.

6.2.3 Our Synthesis:

- POS growth slowdown (FY21–FY23): ~10k tCO_2.
- UPI/QR shift (FY21–FY23): ~20k tCO_2.
- Modeled incremental FY26–FY30: ~45k tCO_2 (guesstimate), reflecting continued UPI growth against FY23 baselines (anchored by RBI Payment System Indicators for overall scale).

Note: Carbon intensity per kWh varies widely by state grid mixes; our incremental projection is conservative and intended for directional planning, not inventory accounting.

6.3 INSTITUTION-LEVEL ROI EXAMPLE (NBFC)

We model a mid-sized NBFC (Rs. 5,000 crore AUM) investing in an AI restructuring stack.

6.3.1 Assumptions:

- Capex (Y1): ₹3 crore; Opex: ₹1 crore/yr
- Benefits: ₹0/₹4/₹6/₹7.5/₹8 crore in Y1–Y5 (time saved, faster TAT, lower write-offs)
- Discount rate: 12%

6.3.2 Results:

- NPV (12%) ≈ ₹10.5 crore, IRR ≈ 100% (illustrative).
- Payback in Year 2 (end).

These numbers are consistent with productivity gains seen in early Indian deployments (e.g., large reductions in CAM preparation time, manual checks, and TAT).

6.3.3 Our Synthesis:

- POS growth slowdown (FY21–FY23): ~10k tCO_2.
- UPI/QR shift (FY21–FY23): ~20k tCO_2.

- Modeled incremental FY26–FY30: ~45k tCO_2 (guesstimate), reflecting continued UPI growth against FY23 baselines (anchored by RBI Payment System Indicators for overall scale).

VII. Sector-wise Indian use cases with quantified payoffs

7.1 Corporate finance (listed/unlisted; MSMEs to large)

1. Invoice-to-Pay (I2P): AI capture + GST validations reduce AP cycle times by 50–70% and defects by 30–60% in pilots; labor + paper savings of ₹8–₹15/invoice are typical in India, plus fewer vendor escalations.
2. Record-to-Report (R2R): Close cycle compressed by 20–40%; LLM-drafted footnotes and variance narratives accelerate audit readiness.
3. Order-to-Cash (O2C): AI remittance advice matching reduces unapplied cash; DSO improvements of 1–3 days are common when linked to dispute triage.

7.2 Banks/NBFCs

- Underwriting/CAM: With AA bank-feeds and GST data, lenders can auto-produce credit packs and sharply reduce "time to sanction"—some Indian deployments report ~85% faster turnaround.
- Monitoring/Early Warning: AI scans GSTR-1/3B consistency, invoice density, and payment rhythms; flagged entities feed risk dashboards.
- Collections: AI-recommended strategies (promises-to-pay likelihood, right channel/time) lower roll rates.
- Reg-reporting: Mapping core banking extracts to RBI returns with built-in explainability addresses model governance under FREE-AI.

7.3 Tax & compliance

- GST: With ~1B+ monthly e-invoices as anchor, AI auto-books, validates ITC, and drafts reconciliations; anomalies trigger vendor follow-ups—saving crores in leakages and penalties, and thousands of FTE-hours per large filer.
- Direct tax: CBDT highlights speed in refunds and updated returns; AI co-pilots can improve prefill accuracy and detect high-risk claims for review.

7.4 Investment research & securities markets

- LLMs screen SEBI filings, investor calls, and media for governance and ESG indicators, with "chain-of-evidence" citations—important under SEBI's proposed AI/ML governance (model audits, human oversight, reporting to SEBI).

7.5 insurance

- **Claims/underwriting:** AI-assisted reading of medical and surveyor reports; automated SE/IRDAI returns using extraction and templates; anti-fraud graph checks.

7.6 Regulators / SupTech

- **Anomaly Detection** on filings (returns, disclosures) and payments; explainable alerts and sandboxed model updates in line with RBI's FREE-AI governance.

VIII. Challenges and how to mitigate

1. **Data quality & OCR noise** (stamps, scans, mixed languages): solve with document layout models, post-OCR correction, dual-pass validation (e-invoice ↔ GSTR ↔ ledger).
2. **Multilingual India**: leverage IndiaAI LMMs and Bhashini-style language stacks for field labels and narrative generation.
3. **Trust, fairness, explainability**: maintain model cards, lineage logs, challenge-response tests; embed FREE-AI and SEBI guardrails—human accountability stays central.
4. **Privacy & security**: AA is consented; use data minimization, on-prem/India-region clouds, and PII tokenization; maintain vendor SLAs and audit rights as SEBI suggests.
5. **Change management & skills**: pair data stewards with domain SMEs; certify reviewers; align incentives to productivity metrics and quality KPIs.

IX. Future Scope (FY26-FY30)

- **LLM-native finance teams**: AI co-pilots drafting Ind-AS notes, board packs, and lender CAMs with citation trails; multilingual output for subsidiaries across India.
- **Real-time reporting:** streaming AA + UPI + e-invoice data into rolling MIS and regulatory dashboards; risk triggers push corrective actions within the same cycle.
- **ESG/CSR analytics:** automated ESG footnotes, spend taxonomy mapping, vendor screening against BRSR indicators; green nudges in payments and procurement.
- **Early warning at population scale:** regulator SupTech fusing payments, GST, and filings to spot emerging risks—aligned to financial stability aims.
- **Domain-specific Indian LMMs:** funded via IndiaAI, enabling better OCR/NLP on local scripts and financial forms.

X. Research Gaps in India vs. Global Work

1. **Ground-truth datasets** for Indian financial forms (publicly shareable, privacy-safe) are limited; **IndiaAI Datasets Platform can** catalyze this.
2. **Model risk benchmarks** specific to FREE-AI and SEBI frameworks are nascent—academia/regulators could publish test suites for fairness, robustness, and explainability in Indian finance.
3. **Carbon accounting for finance ops:** we need standardized factors for paperless flows, branch/ATM energy, and digital infra (GPU) to quantify net impact (mission-critical as IndiaAI compute scales).

XI. Methodology & Reproducibility

- We anchored scale to official sources: RBI Payment System Indicators (July 2025), PIB GST e-invoice (May 2024 figure), SEBI AI consultation listing, IndiaAI Mission outlay/compute, and Sahamati ecosystem stats.

XII. Challenges and How to Mitigate

1. **Pick 2–3 high-yield workflows** (e.g., e-invoice ingestion, bank reconciliation via AA, GST match & file).
2. **Procure responsibly:** insist on FREE-AI/SEBI-ready vendors: data residency, model-risk docs, audit logs, fallback plans, and bias/explainability tests.
3. **Wire into rails:** AA integrations for cash-flow truth; e-invoice/e-way/GSTN APIs; RBI-compliant payment data safeguards.
4. **Design controls:** maker–checker with confidence thresholds and exception queues; retain human accountability for filings and credit decisions.
5. **Track value real-time:** FTE hours released, cycle-time, error rates, write-offs, and tCO_2 proxy (paper avoided, branch/ATM travel avoided, green power share).

XIII. Conclusion

India's financial data fabric—UPI, GST e-invoices, AA, and digital filings—has hit a critical mass at which AI-driven restructuring is no longer optional. With FREE-AI and SEBI guidance, and IndiaAI compute coming online, the next five years can deliver tens of thousands of FTE-years released to higher-value work, ₹ tens of thousands of crores in savings, and meaningful carbon reduction from paperless, branch-lite operations. The key is to move from pilots to governed scale, with transparent assumptions, measurable KPIs, and a bias for explainable automation.

REFERENCES

- **RBI – Payment System Indicators (July 2025):** monthly metrics across UPI, IMPS, cards, ATMs, UPI QR codes.
- **PIB – GST e-invoicing:** "About 110 crore e-invoices were generated during May 2024." (Press Information Bureau)
- **Sahamati (AA):** FY25 disbursements ₹1.6 lakh crore across 1.89 crore loans; ecosystem stats and consents. (The Economic Times)
- **RBI – FREE-AI (Aug 13, 2025):** Committee report; "Seven Sutras" and governance pillars (multiple legal/industry summaries). (Reuters)
- **SEBI – AI/ML Consultation (20 Jun 2025):** listing page and title in SEBI's official repository. (Securities and Exchange Board of India)
- **IndiaAI Mission (PIB, Mar 2024):** ₹10,300+ crore outlay; 10,000+ GPUs compute target; domain LMMs. (Press Information Bureau)
- **PwC India (2024):** Impact of digital payments on GHG emissions in India (UPI/PoS effects). (PwC)
- **ATM energy:** academic/technical estimates ~2.4 MWh/year per ATM (excluding AC). (SCIRP)
- **CBDT Year Ender 2024 (PIB):** faster refunds, updated returns scale; digitization initiatives. (Press Information Bureau)
- **Perfios CAM AI:** press/industry coverage of 85% faster CAM automation. (FF News | Fintech Finance)
- **Clear (ClearTax) + Microsoft:** Azure OpenAI case study (200k+ users).