05

Reimagining invoice validation and streamlining third-party risk management for a hyperscaler

In today's fast-evolving business landscape, organisations are under mounting pressure to optimise financial operations and manage associated risks effectively. Agentic AI, with its advanced capabilities, offers transformative solutions to automate these processes and enable proactive third-party risk management. Through this case study, **Ankit Garg** demonstrates how agentic AI was applied to automate invoice validation workflows for a leading hyperscaler client - enhancing control and compliance. The study also outlines the key benefits and practical challenges encountered during the implementation of the agentic AI solution.

PwC India collaborated with a global hyperscaler to transform the invoice validation process as part of the accounts payable cycle by automating validation checks with agentic AI. The client, that operates across multiple geographies, faced significant challenges in processing thousands of invoices due to its vast and diverse vendor network. Delays in invoice processing, compliance concerns and human errors due to manual processing were identified as key risks. Moreover, conventional systems struggled to keep up with the complexities of their invoice processing operations, especially when working with non-standard invoice formats, varying demands of individual vendors and data files. Lack of consistent control mechanisms across the process was another major challenge.

The client had considered a few alternatives before adopting agentic AI; however, each of them faced the following challenges:

- **Optical character recognition (OCR) and rules-based processing:** This was effective for regular or standard invoices but challenges surfaced while handling almost 40% of the non-standard invoice formats. Varying invoice formats from multiple vendors made it difficult to automate the process to an acceptable level as it became a recurring job and exposed the organisation to huge risks.
- **ML-based solutions:** The solution helped increase the accuracy of the data extraction process but struggled when it came to contextual responses for different styles of invoices or credit notes. Reading and understanding the data and adjusting the context of modified names was a persistent concern.
- **Workflow automation:** Basic automation streamlined the routing of the workflow; however, the client was unable to improve validation accuracy and decision-making.

By deploying agentic AI along with a master orchestration framework, PwC India was able to design a bespoke, scalable automation plan to enhance controls, governance and risk management.

Implementation process

Agentic AI was plugged into the client's invoice processing and validation workflow to enable seamless extraction of data and to automate validation and compliance checks. Key functions of the agentic AI solution included:

- **Data integration:** Consolidation of structured and unstructured data from scanned invoices and multiple complex files.
- **Customised approach:** Using the client's proprietary LLM model with a retrieval-augmented generation (RAG) framework and computational engine helped process a variety of invoice formats by applying complex arithmetic and analytical operations.
- **Compliance mechanisms:** By automating checks, the agentic solution enhanced the compliance of the invoice process with the company's policies. It also assured conformance with multiple vendor contracts to effectively manage third-party risk.
- **Atomic operations and templatised prompts:** Invoice validation logic was broken down into atomic operations such as individual arithmetic operations and rule-based checks where each check was mapped to a templatised prompt. These prompts used structured metadata rather than raw data, enabling consistent and reusable logic across vendors for easier debugging and version control.
- **Agent orchestration:** An agent orchestration layer handled the flow of tasks intelligently and dispatched parts of the invoice to specialised agents depending on context and validation requirements. It enabled inter-validations to be executed in a dependent and sequential order. This was based on their creation sequence with built-in support for fall back and retry mechanisms to tackle failures and anomalies.
- **Modular architecture:** Our solution followed a modular architecture where each component — data ingestion, validation, compliance and reporting — functioned independently and allowed plug-and-play support, parallel processing, fast experiments with new prompts and LLMs or logic changes without needing system-wide changes.
- **Standardised input file templates:** The solution replaced fragile spreadsheet-based macros and traditional automation with scalable, standard templates for better consistency.

Challenges

There were several challenges in the implementation of the solution.

Data quality and integration complexity

The use of agentic AI solutions with existing infrastructure can be challenging since the infrastructure is often traditional and non-interoperable. These environments are often heavily customised for supporting cutting-edge AI, making them difficult to implement with other robots or projects and lacking sufficient computational power.

The client's process was manual and unstandardised with discrepancies in data, misaligned columns and unsegregated data structures. These became an obstacle for generating stable templates for automatic vendor verification.

LLM changes and integration

Since agentic AI technology is constantly evolving, regular updates to LLMs may result in integration difficulties. There may also be processing capability changes with the new versions; therefore, it is necessary to calibrate and train the model repeatedly so that it is compatible with newer versions.

Ethical and compliance concerns

Responsible AI is a critical component of AI adoption. Systems must adhere to certain ethical principles and relevant laws and standards. For instance, key priorities for this project were transparency, accountability and explainability in AI decision-making which are vital factors in minimising compliance risks and fostering stakeholder trust.

Template maintenance and retraining

AI-driven systems must be retrained periodically to keep abreast of new templates, business rules and changes in vendor formats.

RAG limitations and LLM accuracy challenges

The successful use of RAG systems depends on the quality of knowledge the model can use. Misdated or inconsistently labelled data can misguide AI outputs. As a result, some of the output may be incorrect. Therefore, ongoing validation and fine-tuning are necessary to maintain high accuracy and minimise errors in production.

For instance, the LLM used in this project didn't meet some of the validation checks that it met when it processed a data file from the same vendor in a prior iteration.

A retroactive root cause analysis was completed highlighting that the validation template was included along with the data input, causing the LLM to generate a code using its computational capabilities. This led to inconsistent results and ensuring reliable validation was a challenge.

To avoid this, a governance mechanism to manage or control the code produced by the computational engine is required to maintain stable validation of results.

Implementation teams should include the stakeholders early on to decide and design data validation and compliance rules. They should also focus on training teams on AI confidence and on the fallibility to determine which cases require human intervention and review.

The learning curve

The implementation journey of agentic AI came with significant learnings:

01 **Initial over-flagging:** With cautious thresholds and a biased training set, almost 60% of the invoices were flagged in the first two weeks.

Solution: A centralised repository was implemented to maintain a uniform gateway to these vendor-specific validation files.

Key takeaway: AI models require exposure to unstructured, real-world data including edge cases to improve generalisation and reduce false positives.

02 **Vendor-specific patterns matter:** Though accuracy increased to 85% by the end of week three, there was an issue with processing vendor data consistently.

Solution: A standardised, templatised process of prompts with procedural steps to support consistent validations and avoid duplication helped mitigate this.

Key takeaway: In addition to patterns, the AI model also needs to learn vendor context to be reliable.

03 **Managing computational engine variability:** The LLM-generated code executed at runtime led to unpredicted validation outcomes.

Solution: A constrained execution model based only on LLM-generated queries with metadata-driven prompts and a regulatory layer for stability addressed this concern.

Key takeaway: Organisations must refrain from executing LLM-generated codes directly and leverage metadata-driven triggering and controlled environments for consistent results.

04 **Confidence scoring:** One of the key steps of the implementation process was to add confidence levels to all validation output.

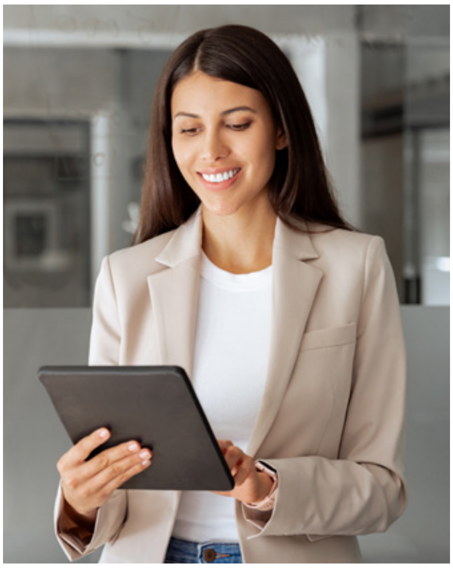
Solution: This approach enabled greater focus on cases with high uncertainty both in terms of insights and the machine-generated output while reducing the human workload.

Key takeaway: Confidence scoring can be easily automated at scale. It also allows human intervention.

Focus points

Implementing agentic AI proved to be challenging due to data inconsistencies, evolving AI models and strict compliance expectations. Addressing these challenges required careful governance, retraining and adaptive design. Some of the factors which hindered the adoption of this technology were:

- **Challenges in achieving end-to-end automation:** The best level of automation hit a ceiling at about 95%. Full automation was impossible due to complicated cases and the level of human judgement needed.
- **Lack of change management:** Though technical deployment had its share of challenges, one of the major hurdles was convincing business users to adapt the new solution.
- **Generic rules:** A one-size-fits-all solution is not valid for inter-department and varied vendor requirements.
- **Uniform vendor treatment:** Model training and standardised templates across vendors were not possible due to the difference in vendor size, scope and customised checks.



Future areas of implementation

Besides invoice validation, some of the areas where the solution can be implemented are:

- **Predictive vendor management:** By processing historical and current data using AI, businesses can predict which vendors are likely to submit incorrect or non-compliant invoices, allowing action to be taken in advance to proactively reduce financial exposure and strengthen relationships with vendors.
- **Dynamic approval of workflows:** Future AI systems will not only be contextually aware but will also utilise dynamic approval hierarchies which can be changed as per real-time business requirements and invoice conditions while still remaining compliant.
- **Cross-process intelligence:** Information obtained from invoice validation can be used in other areas such as purchasing, contract negotiation and vendor performance management which could enable businesses to develop an interconnected IT financial ecosystem.
- **Chat-based interface:** In future, the proliferation of conversational agents will enable finance teams, risk teams and business stakeholders to ask natural language questions about invoice data, risk assessments or compliance statuses to gain access to critical data and make instant, data-driven decisions.
- **New vendor onboarding:** Smartly orchestrated agents and automated workflows could reduce time and cost of new vendor onboarding by automatically creating custom validation templates, prompt configurations and system setups, thereby reducing manual effort and accelerating time-to-value.

Steps for successful implementation

To successfully adopt agentic AI, organisations should:

1. **Start with data readiness** by providing clean and normalised data. This phase would detect and respond to structural inconsistencies such as malformed columns and header mismatch prior to implementation.

2. **Provide centralised access** by creating an easily accessible single source of truth for both structured and unstructured invoice data to track and maintain compliance.

3. **Use metadata over raw data** since raw invoice data in LLMs often yields inconsistent results. Switching to metadata-based prompts ({vendor_type}, {invoice_amount}, {currency}) can deliver consistent output.

4. **Develop a robust design and modular architecture** for the solution with the following steps:

   - **Atomic operation layer:** Break and templatise invoice validations into individual steps based on distinct business rules or calculations.

   - **Agent orchestration:** Route subtasks such as amount checks, vendor authentication or tax validation to specialised agents using a task orchestration layer which considers the context of the request.

   - **Controlled code execution:** Introduce a sandbox for code interpretation like logic or templates that are validated, restricting LLMs from generating and executing dynamic codes at validations.

5. **Templatise prompts** for scalability and faster debug by:

   - **Using prompt templates per vendor/use case:** Leveraging generic prompt templates is effective when applying identical validation checks for multiple vendors. This cuts down the need for re-writing of the check validation template, thus streamlining the process and adding robustness to the system.

   - **Maintaining prompt versioning:** Tracking the evolution of prompts over time, including rollbacks and experimenting, should be part of the implementation process.

   - **Scaling by segment:** Starting with vendors that have similar formats or validation rules before rolling to difficult or high variance sources is critical.

6. **Plan template drift** and model variability by:

   - **Monitoring LLM output variability:** LLMs can produce different answers even when they are given the same input across sessions or releases. Therefore, it is important to implement fallback checks based on deterministic logic to ensure consistency and reliability regardless of changes in LLM behaviour.

   - **Implementing lightweight retraining loops:** Keeping track of validation fails and incorporating re-try in feedback loops with a set limit on the number of retry attempts is essential.

   - **Introducing confidence scoring:** A confidence level should be attached to every validation outcome. This enables the user to scale people-powered automation and detect low-confidence anomalies for reconsideration.

7. **Benefits and measurable outcomes**

   Besides enhanced automation and minimal human intervention, some of the other benefits of the successful implementation of an agentic AI tool are:

   - Improved financial process efficiency and accuracy.

   - Minimal risk of compliance breaches with a proactive audit trail and process in place.

   - Lower operating costs and number of fraud cases make it easier to work with vendors and adapt to new requirements.

   - Automation of data extraction, validation and routing eliminates the need for manual work.

   - A faster and more reliable invoice journey with fewer challenges enables liquidity forecasting for the decision makers.

   - Quick activation of new suppliers with AI-based validation templates which supports special invoice formatting and contractual variations.

Select measurable outcomes of adopting agentic AI solutions for the client were:

Reduced manual processing time
85% decrease in manual processing time per invoice; from an initial spent of 2–3 days down to less than 10 minutes on an average.

Faster deployment
Time from pilot to production deployment was shortened to 1.5 weeks versus the standard 4–6 weeks for traditional automation rollouts.

Lower operational cost
Automation of data extraction, validation and routing processes reduced accounts payable operations costs by up to 60%.

Minimal errors
A massive 85–90% dip in keyed errors meant reduced exposure to compliance penalties and payment disputes.

Standardised templates
The move to 5% standard templates from error-prone macros eliminated the need for extensive maintenance, reducing compliance incidents by 70%.

Centralised invoice repository
The centralised invoice repository was able to reduce compliance issues by 70%, providing a single source of truth regarding the vendors.

Enhanced vendor processes
The organisation was able to simplify onboarding processes and was also able to streamline invoice validation and processing.

Enhanced compliance and audit readiness
The solution helped automate compliance requirements with internal policies to ensure that contract terms and regulatory alignment are fulfilled.

Rapid deployment
The solution was quick to deploy and could go from pilot to a minimum viable product (MVP) in less than 30 days.

Looking ahead

Agentic AI is a game-changing advancement for organisations that can take invoice validation processes from passive legacy automations to a dynamic, intelligent orchestration. Adopting agentic AI offers tangible business benefits including

- quicker resolution of financial tasks
- improved vendor relationship
- higher audit preparedness
- better third-party risk management.

By shifting from rules-based automation to AI-driven orchestration, finance teams can focus on high-value, judgement tasks and rely on AI to process routine validation. Future use cases of agentic AI could include predictive vendor profiling, dynamic approval processes and conversational analytics; however, the focus of the C-suite and finance function should be on developing collaborative human-AI teams where AI is not a mere tool but an adaptive co-pilot that helps create smarter, faster and more resilient financial ecosystems.

Contributing to this article were **Anirudh Singh Rana**, **Sachin Parashar** and **Rodney D'Souza.**