



# **The finance leader's guide to evaluating agentic AI**

Most finance AI vendors claim agentic capability. Few can demonstrate it. This guide gives finance leaders a practical framework for evaluating agentic AI platforms. It explains what real agentic capability looks like, presents questions to expose vendors who can't deliver it, and identifies the red flags that separate genuine agentic platforms from automation relabeled as AI so you can easily compare vendors.

## What is agentic AI?

Agentic AI is a system of AI agents that perceive their environment, reason through complex decisions, and execute tasks across interconnected systems without requiring human instruction at each step. Three AI technologies work together in a well-built agentic finance system: machine learning models that recognize patterns and anomalies, large language models that interpret unstructured documents and policy language, and orchestration layers that coordinate the completion of multi-step processes by multiple agents.

A contextual reasoning layer enables the system to understand nuance as it extracts data. It builds a library of patterns and applies those to its predictions as it creates new data points. Unlike general-purpose AI models, specialized finance AI agents also include verification layers, finance-specific training, and human-in-the-loop (HITL) oversight.

Where generative AI is assistive, sharing helpful suggestions and waiting patiently for human approval, agentic AI is authoritative, with the ability to make plans and carry out tasks independently. The distinction between authoritative and assistive AI determines whether your investment reduces manual workloads or simply adds a smarter interface to the same manual process.

“Back in the day, 40 accountants spent approximately 5 hours per month auditing expense reports. That is a combined 200 hours auditing these expense reports. Now it's 1 accountant spending maybe 2 hours a month.”

**Finance Leader**  
Flight Centre Canada

**FLIGHT  
CENTRE®**

## How does agentic AI benefit finance operations?

In finance, an AI agent can receive an invoice, verify vendor data against your enterprise resource planning (ERP) or a receipt against your expense management system (EMS), check transactions against policies, flag anomalies, route for approval, reply to emails regarding status, and complete the full workflow autonomously. The agent handles exceptions by reasoning through context, the way an experienced Travel & Expense (T&E) auditor or accounts payable (AP) team member would.

The result is a system that acts on the team's behalf with speed, accuracy, and the institutional knowledge necessary to scale with your business.

### What AI agents do across finance operations

AI agents handle end-to-end workflows across every major area of finance operations. Organizations that deploy agentic AI across these functions report measurable gains: faster cycle times, higher audit coverage, and the ability to scale finance operations at higher transaction volumes without adding headcount.

**Accounts Payable (AP)** Agents process invoices end-to-end, handle exceptions, manage vendor communications, and prevent duplicate payments without manual queuing.

**Travel & Expense (T&E)** Agents audit every transaction against policy, flag out-of-policy spend, and resolve exceptions without requiring a reviewer to work a backlog.

**Compliance** Agents continuously monitor transactions for potential fraud, policy violations, and regulatory risk, taking action before exposure escalates.

**Financial Planning & Analysis (FP&A)** Agents surface relevant data anomalies and trends that inform faster, more accurate decisions.

## Why is finance AI adoption accelerating now?

Late 2022 marked the turning point when generative AI became widely available, expanding the functions finance could automate. Today, agentic AI extends that capability with autonomous action, completing workflows, handling exceptions, and learning from every transaction without IT involvement.

The structured, rule-based nature of finance processes makes them well-suited for agentic AI. Leaders of Finance, Shared Services Operations (SSO), and Global Business Services (GBS) are deploying agentic AI to process expense reports, audit invoices, and flag compliance risks in real-time. With it, they're driving real business growth by running leaner teams while processing more transactions with greater accuracy.

That accuracy is fundamental; incorrect actions carried out autonomously create real financial and reputational risk. Data privacy, trust, and security around proprietary and financial data are legitimate concerns for finance teams. With those concerns, global economic volatility has increased the competitive pressure to adopt AI and do more with less.

# What makes a finance AI system truly autonomous?

A finance AI system is considered autonomous when it can complete any business process without requiring human action at each step. Use these four defining capabilities as your evaluation rubric.

## 1. What intelligent action does the AI take on its own?

A genuinely agentic AI learns from your review process to make accurate, contextual decisions without assistance. It completes finance workflows end-to-end, offers preemptive recommendations, and coaches auditors and managers when intervention is needed. AI agents go beyond answering questions; they execute. From drafting vendor communications to processing expense reports, they act based on reasoning rather than rigid rules. Over time, the system becomes an active participant in your finance operations, not a tool your team has to manage at every step.

If a vendor's AI produces outputs that humans must act on to complete every workflow, it is a decision-support tool, not agentic AI.

## 2. How does the AI understand the context of the data?

What separates a genuinely agentic system is the ability to understand the meaning of the words within the context of a document. Look for an AI solution that evaluates risk based on context clues read across an entire document, not just individual fields. Can it accurately categorize a personal care item on an expense report, even if it has never seen that exact item or brand? Can it flag anomalies indicating a suspicious vendor relationship based on patterns across multiple invoices, without a direct rule match?

Ask vendors to demonstrate this specifically with documents that look like yours.

## 3. How quickly can the AI adapt to new policies and formats?

One of AI's greatest advantages is its ability to access "dark data," the historical transaction data that finance teams cannot easily research or use themselves. Agentic AI uses this data to build a library of recurring patterns, predicting the completion of unfamiliar line items and recognizing document formats it has never processed. It applies learned patterns to determine the most appropriate action to take. Your team retains 100% of the institutional knowledge the AI has accumulated, including processes, policies, and document formats, regardless of staff turnover or business changes.

Confirm that the vendor that the AI can adapt without IT involvement when your business scales or policies change.

## 4. How does the AI know when to act and when to escalate?

AI cannot replicate all human expertise and judgment. Agentic AI should only take action when its confidence is high. It should be transparent about when it is not. The system must acknowledge low-confidence predictions and ask for user feedback, handing off the decision to a human team member with the right expertise. It should always produce a clear audit trail of every action it took, and why. This allows finance teams to verify, learn from, and trust the autonomous decisions of their governed hybrid workforce.

These four components separate true agentic systems from decision-support tools that still require humans to act on every AI output.

# What questions should you ask a finance AI vendor?

These questions map directly to the four capability components above. Apply this framework to every vendor you evaluate. Look for those who can outline specifically how their systems handle each dimension.

## How autonomous is the AI and what governance exists?

This is the most important category and the one where vendors' answers vary most. Agentic AI takes autonomous action in your financial systems. You need to know exactly what it can and cannot do on its own, and how its safeguards prevent errors from cascading before you see them.

- What does your AI agentic architecture look like, and how do multiple agents coordinate on a single workflow?
- What actions can your AI agents take autonomously, and what always triggers human review?
- How does the system prevent incorrect autonomous actions or cascading errors?
- What audit trail exists for every decision and action the agent takes?
- How does the human-in-the-loop model work, and can we configure autonomy thresholds?

## How does the system protect sensitive financial data?

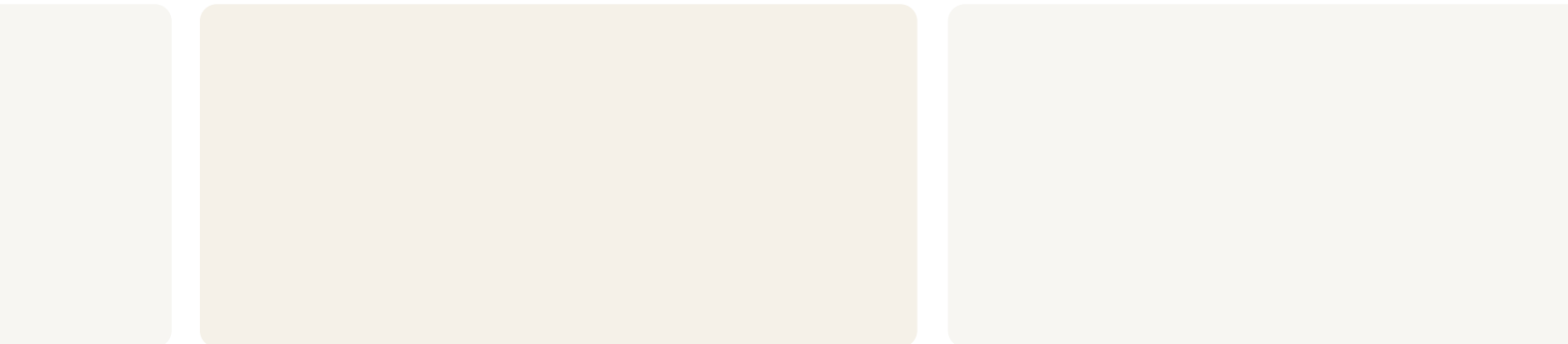
To safeguard sensitive financial information, agentic AI solutions must be designed with robust data privacy frameworks. The AI architecture should anonymize and secure your data so no proprietary or confidential information is ever publicly exposed. Every agent action should be logged and traceable.

- What data privacy certifications and compliance standards does your solution meet?
- How does your AI anonymize and secure our proprietary information?
- How does the system maintain the integrity and confidentiality of third-party data?
- How are agent actions logged, and who has access to that audit trail?

## How does the AI adapt to our policies and systems?

AI agents must be flexible enough to align with company-specific rules, policies, and regulations. You should be able to configure the AI to reflect your unique operational standards, compliance requirements, and existing corporate frameworks, without requiring IT involvement for every adjustment.

- How quickly do your AI agents learn new policies, document formats, or invoice workflows?
- What does that learning process look like in practice?
- Can agents rapidly scale up or down as our business changes?
- How are agents updated when our policies change, and who is responsible for that process?



### **Can you validate your autonomous capabilities before purchase?**

Request a live demonstration of autonomous capabilities. Combine this with a review of testimonials and performance data from existing customers. Pay close attention to outcome data from customers past the one-year mark. Long-term performance tells you more than early adoption numbers.

- Can you present a live demo of your AI agents completing a finance workflow end-to-end, without human intervention?
- Can you show me how the system handles a document or exception it has never encountered before?
- Can you share references, testimonials, and outcome data from customers who have used your solution for a year or more?

### **How accurate is the AI and how does it handle errors?**

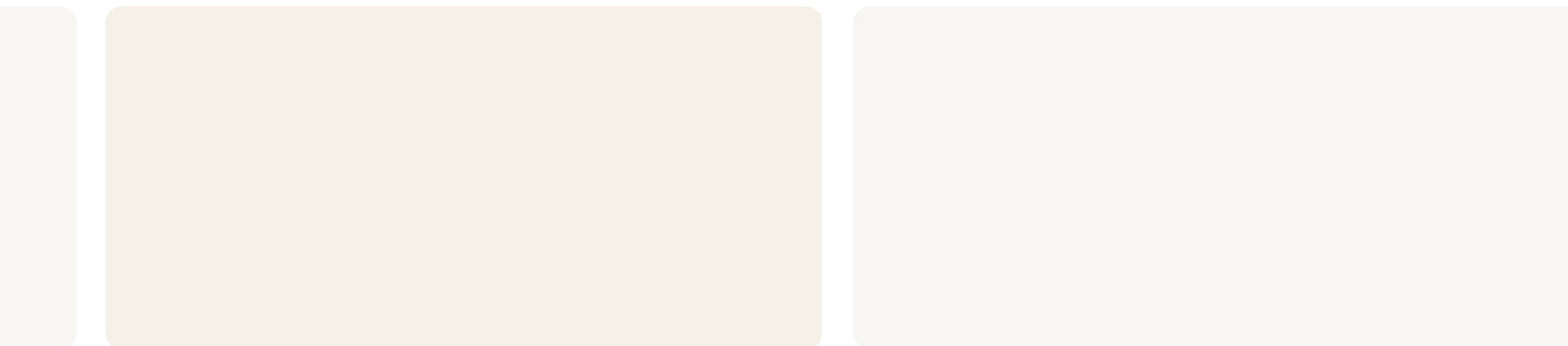
In finance, accuracy is non-negotiable. Autonomous actions amplify both accuracy and error, so choose an AI that finance domain experts have trained deeply and fine-tuned. The source of training data is critical to minimizing bias and ensuring accuracy across industries and geographies.

- Has your AI been fine-tuned with deep domain expertise in finance?
- What is the source of your training data, and does it represent various industries and languages?
- What safeguards prevent hallucinations and other AI-generated errors from triggering incorrect autonomous actions?
- How does the system communicate uncertainty before it acts?

### **How does the AI integrate with existing financial systems?**

AI agents must integrate with your existing financial systems to act within them. This is also where change management plans separate providers who deliver outcomes from those who merely deploy software.

- How do your agents integrate with our ERP, EMS and/or other financial systems?
- What IT infrastructure is required, and how will IT be involved?
- How quickly can we expect a return on our investment?
- Do you offer ongoing training and support throughout the deployment lifecycle?



# 10 warning signs a vendor's finance AI claims are overstated

Many finance AI vendors will claim agentic capability. The difference between genuine autonomous AI and AI-labeled automation is specific, and vendors who can't demonstrate the specifics are usually hiding the gaps. These ten red flags are the most reliable signals that a vendor's claims do not match their system's abilities.

## Red flags to watch out for

1. Vague descriptions of what the AI actually does autonomously versus what requires human action
2. An inability to explain how their AI agents are orchestrated, what decisions they make autonomously, and what safeguards prevent errors from cascading
3. A refusal to provide a live demonstration of an autonomous, end-to-end workflow using documents similar to yours
4. Emphasis on human-led review for all agent outputs; if the AI advises but humans must act on every decision, it's a decision-support tool, not an agentic system
5. A lack of specific use cases or customer testimonials with measurable results
6. You would be an early adopter of their agentic technology; your production environment is their testing environment
7. The company has a small team and a limited operational history of building AI for finance; agentic AI for enterprise finance requires deep domain expertise accumulated over years
8. AI is "tacked onto" existing automation technology, such as a solution that uses AI to speed up optical character recognition (OCR) data extraction but has no document understanding or autonomous action capability
9. The solution has a limited ability to learn and adapt to new suppliers, document formats, or processing requirements; any adjustment requires IT specialists
10. There is no clear audit trail for agent decisions and autonomous actions, making it impossible to verify why the AI took a specific action, identify where something went wrong, or defend its decisions to auditors

## How AppZen can help

AppZen has been building AI-enabled finance products for over a decade. In that time, our Expense Audit, Card Audit, and Autonomous AP, have processed millions of transactions annually for global enterprises. Our technology stack combines proprietary machine learning models, finance-trained large language models, and AI Agents that operate as the last-mile processing layer, creating truly autonomous finance operations.

With AppZen AI Agent Studio, your finance team can build, deploy, and manage AI Agents that connect to your existing systems and complete finance workflows without extensive manual work. Our agentic AI platform orchestrates how those Agents work together, routing tasks, resolving exceptions, and escalating decisions when human judgment is required.

Our proprietary technology adjusts to customers' unique business needs, without requiring IT for every change, and continuously learns from every document, policy update, and exception it handles. We are constantly building new AI Agents, features, and products designed to future-proof our clients' finance operations as their businesses evolve.



AppZen is the leader in autonomous spend and compliance management, revolutionizing enterprise finance operations with its proprietary AI platform. Built specifically for finance, AppZen's solutions accurately automate expense audits, card audits, and invoice processing. They integrate seamlessly into your existing processes and learn from your workflows to drive operational efficiencies, ensure compliance, and help your business grow. Empower your AP, T&E, and Card teams to transform your business with AI. Learn why the world's largest companies trust AppZen's AI platform to automate their finance operations. Visit [www.appzen.com](https://www.appzen.com) today.

### Related resources

[The intelligent finance office: A CFO's blueprint for AI-driven success, AppZen.](#)

[Agentic AI for AP: A new era for finance leaders, AppZen.](#)

[The CIO's guide to modernizing the finance tech stack with AI, AppZen.](#)

[Eight essential AI capabilities for the next generation of Shared Services leaders, AppZen.](#)

[AI Agents are reversing 20 years of outsourcing strategy, AppZen.](#)

[Why finance leaders are choosing agentic AI over spend monitoring, AppZen.](#)

# Glossary

Here are some important terms to know when evaluating agentic AI solutions for finance.

| Term                                 | Definition                                                                                                                                                                                                                                                                                                                                      |
|--------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>2-way match</b>                   | A 2-way match in accounting is a process used to match purchase orders (POs) and invoices when validating transactions.                                                                                                                                                                                                                         |
| <b>3-way match</b>                   | A 3-way match adds a layer of verification to a 2-way match before processing a payment by cross-checking the purchase order (PO), invoice, and receiving report against one another.                                                                                                                                                           |
| <b>Agentic AI</b>                    | AI systems that are designed to take autonomous action in pursuit of defined goals are considered agentic. Unlike AI systems that generate outputs for humans to act upon, agentic AI reasons through problems, makes decisions, and executes tasks across connected systems, without waiting for human instruction at each step.               |
| <b>AI agents</b>                     | Agents are AI programs designed to perform tasks autonomously. They understand the context of what they read, reason through decisions, and take action to achieve specific objectives without human intervention at each step. In finance, AI agents can independently complete full workflows, from document processing to payment execution. |
| <b>AI agent orchestration</b>        | An orchestration layer coordinates multiple AI agents working together to complete complex, multi-step processes. Orchestration involves routing tasks to the right agents, managing dependencies, and determining when human review is required.                                                                                               |
| <b>Algorithms</b>                    | Algorithms are mathematical rules or instructions given to an AI system to help it learn from data and make decisions. These are the fundamental building blocks of any AI program.                                                                                                                                                             |
| <b>Artificial intelligence (AI)</b>  | Artificial intelligence, or AI, involves computer systems that perform tasks typically requiring human-level intelligence, including reasoning, interpreting natural language, and recognizing patterns in data.                                                                                                                                |
| <b>Automation</b>                    | Automation refers to the use of technology to perform tasks without human intervention. Automated systems are not context-aware and lack reasoning capabilities, which distinguishes them from AI agents. Examples of automated finance tasks include structured data entry or report generation.                                               |
| <b>Autonomous finance operations</b> | Autonomous finance operations is the application of context-aware AI technologies, like AI agents, to fully automate complex, multi-step finance processes accurately and efficiently without human oversight.                                                                                                                                  |
| <b>Autonomous systems</b>            | Systems are considered autonomous when they are capable of performing tasks or making decisions independently from human input. Autonomous systems enhance efficiency and speed decision-making using feedback loops, giving them the ability to learn from new information and adapt to changing environments.                                 |

| Term                                          | Definition                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|-----------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Bot / Robotic Process Automation (RPA)</b> | Robotic Process Automation, sometimes referred to as “bots,” are software applications programmed to perform repetitive tasks automatically, reducing errors and the need for human labor. Unlike AI agents, RPA follows fixed rules and cannot reason through exceptions or adapt to new inputs without reprogramming.                                                                                                                                     |
| <b>Computer vision</b>                        | Computer vision is a field of AI that enables computers and systems to derive meaningful information from digital images, videos, and other visual inputs and act based on that information.                                                                                                                                                                                                                                                                |
| <b>Deep learning</b>                          | Deep learning is an advanced form of machine learning involving neural networks, inspired by the human brain. These neural networks are made up of multiple layers of interconnected nodes of mathematical equations, with feedback loops that improve the accuracy of the system over millions of repetitions. Deep learning is particularly useful for understanding incredibly complex patterns, such as those involved in image and speech recognition. |
| <b>Electronic Data Interchange (EDI)</b>      | Electronic Data Interchange is the automated, computer-to-computer transfer of business documents (like invoices) between different companies, using a standardized, highly structured format. This paperless processing improves the speed, accuracy, and efficiency of data exchange for documents such as purchase orders.                                                                                                                               |
| <b>Enterprise resource planning (ERP)</b>     | Enterprise resource planning systems integrate core business processes, such as inventory, order management, accounting, HR, and CRM, into a single platform. They simplify operations and the speed the transfer of information across an organization.                                                                                                                                                                                                    |
| <b>Expense management system (EMS)</b>        | Expense management systems automate the submission, approval, and reimbursement of employee-incurred expenses, helping to reduce operator error and enforce policy compliance.                                                                                                                                                                                                                                                                              |
| <b>Finance AI</b>                             | Finance AI refers to artificial intelligence technologies, including machine learning, generative AI, and agentic AI, used in financial applications to improve decision-making, automate workflows, and reduce risk.                                                                                                                                                                                                                                       |
| <b>Generative AI (GenAI)</b>                  | Generative AI is AI capable of generating new content based on its training data. In finance, GenAI enables systems to interpret unstructured documents, draft responses, and communicate with users in natural language. It is a foundational technology layer for agentic AI.                                                                                                                                                                             |
| <b>Human-in-the-loop (HITL)</b>               | When an AI system is designed to involve a human decision-maker at specific points in an automated workflow, this is referred to as human-in-the-loop design. HITL thresholds define which decisions an agent handles autonomously and which are escalated to human reviewers.                                                                                                                                                                              |
| <b>Large language model (LLM)</b>             | Large language models are types of artificial neural networks trained to understand and generate human-like text. In agentic AI systems, LLMs enable agents to interpret unstructured financial documents, policy language, and vendor communications.                                                                                                                                                                                                      |

| Term                                       | Definition                                                                                                                                                                                                                                                                                                   |
|--------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Machine learning (ML)</b>               | Machine learning is a subset of AI that enables computers to learn from data and improve performance over time without requiring explicit programming. ML serves as a foundational layer in agentic AI systems.                                                                                              |
| <b>Natural language processing (NLP)</b>   | Natural language processing is a branch of AI and computer science that enables computers to understand, interpret, and generate written and spoken human language. NLP allows people and computers to interact more intuitively.                                                                            |
| <b>Neural network</b>                      | A neural network is a set of machine learning algorithms modeled after the human brain. It consists of layers of interconnected nodes, called neurons, that process data, learn to recognize patterns, and can make decisions without needing explicit, hard-coded rules.                                    |
| <b>Optical character recognition (OCR)</b> | OCR is the technology used to convert documents, such as scanned papers, PDFs, or images, into editable and searchable data.                                                                                                                                                                                 |
| <b>Policy compliance</b>                   | Policy compliance refers to people adhering to the rules set down by an organization to govern their actions and avoid risks, ensuring that operations are conducted ethically and legally.                                                                                                                  |
| <b>Politically exposed person (PEP)</b>    | A PEP is an individual who is or has been entrusted with prominent public functions, and includes their close associates and family members. Financial institutions monitor PEP transactions closely to prevent corruption and money laundering.                                                             |
| <b>Standard operating procedure (SOP)</b>  | A standard operating procedure is a set of step-by-step instructions compiled by an organization to help workers carry out routine operations. Its goal is to achieve efficiency, quality, and uniformity of performance while reducing miscommunication and improving compliance with industry regulations. |
| <b>Structured data</b>                     | Structured data is any data that resides in a fixed field within a record or file, such as data contained in relational databases and spreadsheets.                                                                                                                                                          |
| <b>Touchless processing</b>                | Touchless processing refers to the automation of business process workflows to minimize or eliminate the need for human intervention. In agentic systems, AI agents that handle the full transaction lifecycle, including exceptions, without manual work at every step, are considered touchless.           |
| <b>Transformers</b>                        | Transformers are a type of AI architecture that processes sequences of data, like words in a sentence, by analyzing how each element relates to the others. Large language models used in generative AI and agentic AI systems, like GPT and Claude, are built on transformer architecture.                  |
| <b>Unstructured data</b>                   | Unstructured data is information that does not have a pre-defined data model or organization. This includes documents, images, audio, and video files. Agentic AI's ability to reason over unstructured data is a key differentiator from earlier automation technologies.                                   |

| Term            | Definition                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|-----------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Vectors</b>  | <p>Vectors are numerical representations of data that allow machine learning models to process and compare information mathematically. In AI language applications, words and concepts are converted into vectors and “mapped,” so models can identify patterns and relationships within text. For example, models can recognize that “dog” and “puppy” are closely related. Context is important in determining meaning; the vector for “bank” changes depending on whether the surrounding text is about finance or a riverbank.</p> |
| <b>Workflow</b> | <p>A workflow involves the step-by-step sequencing of processes through which a transaction document, such as an expense report or an invoice, passes from initiation to completion. Workflows organize and simplify routine finance operation processes.</p>                                                                                                                                                                                                                                                                          |