inbenta

CLOSING THE PROMISE GAP:

How Encore Supports the OCC Supervisory Framework for AI in Banking

Executive summary

Banks can use AI in customer service with confidence when every answer comes from content the bank has already approved. That is the core of **Encore**, and it fits closely with what the OCC now asks examiners to look for.

In November 2026, the Office of the Comptroller of the Currency (OCC) released its Supervisory Framework on Bank Use of Super Intelligence. Under a recent executive order, "Super Intelligence" (SI) is the government's new term for what we have called artificial intelligence. The framework gives examiners sixteen components to review, grouped into six pillars, along with an illustrative four-tier risk model.

The framework does not create a separate AI risk category. Instead, it asks how AI changes risks banks already manage, such as compliance, operational, and customer risk.



Central principle is simple

Controls should grow with the consequences of failure, not with how advanced the technology is

Customer-facing generative AI raises some of the framework's hardest questions. These include hallucinations about fees, rates, and policies; answers that change from one ask to the next; and outputs that cannot be traced to a source. **Encore** was built to resolve those questions by design.

01

Before go-live

Large language models do their work before go-live, to ingest, organize, and prepare content.

02

Experts approve

Internal experts who have appropriate expertise and authority approve that content.

03

Customers get answers

Customers then receive responses drawn from human reviewed content, and every answer can be traced back to its source.

This paper explains what the framework asks of banks, where customer-facing AI is most exposed, and how **Encore** helps banks meet those expectations. It also explains the parts that always remain the bank's responsibility, because no vendor can carry those alone.

What the OCC framework asks of banks

The framework asks banks to know what each AI system is allowed to do, what happens if it is wrong, and how they would know. Examiners are told to apply judgment based on each bank's size, complexity, and risk profile. The OCC calls it a living document that will change as technology changes.

The sixteen components are grouped into six pillars:



Purpose

A clear, testable statement of what the system may and may not do



Risk

An assessment of what happens if an output or action is not fit for purpose



Control

Human oversight, model risk management, data governance, privacy, compliance, and cybersecurity



Oversight

Who monitors the system and who is accountable for it



Resilience

The ability to detect, contain, and recover from failures while the business keeps running



Assurance

Testing, documentation, vendor risk management, and independent audit

The framework closes with four illustrative risk tiers.

TIER 1: CRITICAL

Covers direct, hard-to-reverse actions with high impact, such as an autonomous payment agent

TIER 2: HIGH

Includes agent-initiated customer communications

TIER 3: MODERATE

Covers decision support with a human in the loop

TIER 4: LOW

Covers internal tools with no customer or financial data exposure

The applicable risk tier depends on consequence, reversibility, autonomy, data sensitivity, explainability, and provenance. As the OCC notes, a simple agent that can move money can be riskier than a sophisticated model that only drafts internal documents.

The hard problems for customer-facing AI

Four concerns appear repeatedly in the framework, and each is sharpest when AI talks directly to customers.



Hallucinations

The OCC warns that hallucinations about fees, rates, account status, or policies can create legal exposure and customer harm. A fluent but wrong answer about a fee is still a wrong answer, and the customer has no way to tell.



Non-determinism

Generative AI can give different answers to the same question. The framework says repeated testing and consistency thresholds may be needed, because a single successful test run proves little.



No verified source

The OCC notes that many generative outputs have no verified source to test against. That makes it hard to say whether an answer is right, and hard to measure how often it is wrong.



Traceability

Examiners want banks to reconstruct material outputs: the inputs, the sources used, the approvals, and the result. When a model writes each answer fresh, that trail is often thin.

These are not reasons to avoid AI

They are reasons to choose an architecture that answers these questions before an examiner asks them.

How Encore works

Encore is a knowledge-first, glass-box AI platform for customer experience. It uses large language models (LLMs) where they are strongest and keeps them out of the place where risk is highest:
writing live answers to customers.

Content is prepared up front

Before go-live, LLMs ingest, process, tag, and embed the bank's content, and help create the answers

Internal experts approve it

Answers are built from the bank's own knowledge, not generated, then matched to questions and approved by the internal experts that own the content before customers ever see it. The answer exists before the question is asked

Customers get approved answers

In production, Encore does not use generative AI to write answers on the fly.
Customers receive answers drawn from the approved knowledge base

Every answer is traceable

Because answers come from known content, each one can be traced back to its source.
That is what we mean by glass box

Gaps are surfaced, not guessed

When a customer asks something that the knowledge base cannot support, an internal component of Encore, called Elevate, flags the gap so an experienced staff member can decide how to fill it. Elevate also helps improve answers and automate maintenance over time.



Mapping Encore to the framework

Encore speaks most directly to the components where customer-facing AI carries the most risk.

What examiners look for	OCC component	How Encore helps
A testable boundary; activity outside it treated as a control failure	**2.1** Defined Purpose and Permitted Use	Encore answers only from approved content. Questions outside that scope are flagged by Elevate, not answered with a guess.
Meaningful review by internal staff with real authority	**2.3** Human-in-the-Loop	Internal subject matter experts review and approve content before go-live, so human judgment shapes every answer customers receive.
Handling non-determinism and reproducibility	**2.5** Model Risk Management Extensions	The same question leads back to the same approved answer, which makes repeat testing straightforward.
Control of hallucinations about fees, rates, account status, and policies	**2.8** Compliance	Customer answers are not generated on the fly, so representations come from content the bank approved.
Ongoing measures of output quality and drift	**2.12** Performance Monitoring	Elevate shows where the knowledge falls short and supports regular improvement.
Testing that works even without a single expected result	**2.13** Testing	The approved knowledge base gives the bank a verified source to test answers against.
Ability to reconstruct material outputs	**2.14** Documentation & Auditability	Glass box design lets each answer be traced back to its source content.
Controls that scale with consequence and autonomy	**Section 3** Risk Tiering	Encore's Q&A, Workflows, and Agents levels let a bank grow autonomy step by step.

The common thread is that **Encore** moves the hardest questions to before go-live, where subject matter experts (SMEs) can review the work, rather than leaving them to the moment a customer asks.

Governance: front-line teams and IT working together

With Encore, IT is always part of the picture. But, importantly, the project kick-off is not conditioned on an IT team that likely has its own set of competing priorities.

The internal experts closest to customers know the questions customers ask and the answers they need. Encore lets those front-line teams begin building and shaping an assistant without waiting on developers. That speeds up the work that most depends on business knowledge.

IT, security, risk, and compliance teams still play their full roles. They govern integration, access, security review, and approval before anything reaches production. This matches the OCC's call for clear ownership and approval authority at meaningful stages of the lifecycle (component 2.4).

The result is a shared model of accountability

The business owns what the assistant says. Technology and risk teams own how it is connected, secured, and approved. Both are needed, and neither has to wait on the other to begin.



A tiered path: Q&A, Workflows, Agents

Encore's three levels let a bank match autonomy to its controls, which is exactly how the OCC asks banks to think about risk.

Q&A

WHAT IT DOES

Answers customer questions from approved content

HOW RISK CHANGES

Informational, reversible, no direct action on accounts

CONTROLS TO EMPHASIZE

Content approval, Elevate gap review, answer testing

Workflows

WHAT IT DOES

Guides customers through defined, bank-designed steps

HOW RISK CHANGES

More structure and data involved; steps are set in advance

CONTROLS TO EMPHASIZE

Process approval, data access limits, step-level logging

Agents

WHAT IT DOES

Takes actions on the customer's behalf

HOW RISK CHANGES

Direct action raises consequence and reduces reversibility

CONTROLS TO EMPHASIZE

Tight permissions, human oversight,
kill switches, incident plans

Each bank classifies its own use cases, so final tiers will vary.
Still, the pattern is clear. A bank can start with Q&A, prove its controls, and then add workflows and agents as its governance matures.

Shared responsibility

The OCC is clear that accountability stays with the bank, even when a vendor builds or hosts the technology (component 2.4). No platform can meet the framework on a bank's behalf, and we would not want to suggest otherwise.

What remains with the bank includes its risk appetite, use-case inventory, risk classification, independent validation, internal audit, and board reporting. What a good partner can do is make each of those easier, with clear answers and evidence the bank can use.

The framework's guidance on third-party risk (component 2.9) suggests questions every bank should ask any AI vendor, including Inbenta:

01.	**Which models are used, where, and how will you notify us when they change?**
02.	**Is our data ever used to train or improve models offered to other customers?**
03.	**Where is our data stored and processed, and how long are prompts and logs kept?**
04.	**How do you define an AI incident, such as a hallucination or misfire, and how quickly will you tell us?**
05.	**What protections are in place against prompt injection and other AI-specific attacks?**
06.	**How can we pause, contain, or roll back the system if something goes wrong?**
07.	**How are responsibility and liability assigned for errors that cause customer harm?**

Inbenta welcomes these questions

Inbenta is regularly addressing queries like these from customers that are subject to existing AI laws and regulations that are already in place, such as the EU AI Act.

Conclusion

The OCC framework does not ask banks to slow down. It asks them to know what their AI is doing, prove it works, and stay accountable for the results.

For customer-facing AI, that starts with the question: where does each answer come from?
With **Encore**, the answer is content the bank has reviewed and approved, with a clear trail back to its source.
Front-line teams can begin the work, IT and risk teams govern it, and customers get answers they can trust.

That is how we close the Promise Gap: AI that delivers on what it promises, in a way examiners can see.

This paper summarizes the OCC Supervisory Framework on Bank Use of Super Intelligence (November 2026).
It is for general information and is not legal or regulatory advice.

Read the full framework: **OCC Bulletin 2026-54**

inbenta

Put Encore through your vendor review

We've answered the seven questions above in writing, ready for your third-party risk file.
Book 30 minutes and we'll walk your risk and IT leads through them.

Schedule a call