GUALITY TIMES

DRIVING GENAI SUCCESS IN FINANCIAL INSTITUTIONS THROUGH QUALITY ENGINEERING



Experience & Digital Engineering | Data & GenAl | Cloud | Quality Engineering

WELCOME

To adapt in today's rapidly changing technology landscape, quality engineering (QE) must transform. The question is, "how?"

The Quality Times is our attempt to answer that question. Our contributors belong to an exclusive community of advanced QE leaders, regularly exchanging insights on the industry's latest trends and technologies. Having guided their organizations through waves of disruption, they bring unique perspectives to help you navigate your own modernization journey.

The potential impact is enormous. Through their experiences, they make a compelling case for QE investment, showing how to capitalize on new technologies and increased digitalization while delivering secure, reliable and captivating user experiences.

Our first article features Aaron Haehn, Director of Quality Engineering at T-Mobile, who brings deep expertise in modernizing QE practices and implementing GenAl solutions, alongside series co-author Dror Avrilingi, Amdocs' VP Head of QE, Data & GenAl Studios. With hundreds of successful QE and GenAl transformations across financial services, media, and telecommunications globally, Dror enriches Aaron's insights with deep cross-industry perspective. The article includes insights from our comprehensive study: a survey of 200 financial enterprises across Americas, EMEA and APAC, plus interviews with leading QE professionals.

These experiences with GenAl reflect a broader industry shift. The technology is fundamentally changing how organizations imagine the future of work. For quality engineering, GenAl has a two-pronged impact: it changes how quality engineers test and needs to be tested itself.

Looking ahead, as GenAl automates more of the software development process, quality engineering will become even more critical. Organizations will need skilled professionals to orchestrate and validate these systems, ensuring software meets the highest standards of quality, reliability, and security.

As these changes unfold, we hope you'll find this series thought-provoking, engaging and, above all, useful, as you think about how your organization's QE function needs to adapt to the remarkable pace of technological change. Together with our dedicated partners and passionate evangelists, we aim to shape the future of quality engineering.





Aaron Haehn Unleashing the Power of GenAl in Quality Engineering

As the Head of Quality Services at T-Mobile USA, Aaron is immersed in the GenAl transformation. Aaron's 30 years of IT experience range from leading the development of transformative strategies to implementation across all aspects of the IT ecosystem. His experience spans operations, monitoring, system reliability, systems/software development, and IT problem solving. Aaron believes in the power of collaboration with partners to redefine quality – the goal should be to make it a part of an organization's ethos, rather than something that they do. By fostering a culture of excellence, Aaron aims to elevate the practice of quality to new, sustainable levels of success and simplicity.



Dror Avrilingi Architect of the NextGen QE Consortium

As Head of QE, Data & Al Studios at Amdocs, Dror brings over 25 years of quality engineering experience and is currently driving the rapid expansion of AQE's GenAl use cases. Previously, as Head of Amdocs Quality Engineering (AQE) International, he oversaw global strategy and led significant growth. Before that, as CTO and Head of Strategy for AQE, he established innovative "new ways of working," led a workforce upskilling program, and managed the R&D team, resulting in 13 patents. Passionate about advancing the frontiers of QE, Dror founded the NextGen QE consortium to discuss cutting-edge trends, and aims to revolutionize the field with his fellow evangelists.





THE SMART MERGER OF THE DEVELOPER AND THE QUALITY ENGINEER



GENERATIVE AI'S IMPACT ON QE



QE'S IMPACT ON GENERATIVE AI



HOW TO INJECT GENERATIVE AI INTO QE





HOW QE MAXIMIZES BUSINESS BENEFITS FROM GENERATIVE AI



CONCLUSION

THE SMART MERGER OF THE DEVELOPER **AND THE QUALITY** ENGINEER

THE SMART MERGER OF THE DEVELOPER AND THE QUALITY ENGINEER

inancial institutions are transforming their software development using

Generative AI (GenAI), significantly expanding their ability to provide increasingly sophisticated digital service solutions to their customers. But these advances carry with them a certain amount of risk. In banking, quality cannot be compromised. While GenAl can boost developer productivity by 20-50% and help move customer-focused services to market faster, banks still need to ensure that customers retain the highest level of trust in the banks' systems. Banks still need to protect assets and information, and adhere to strict regulatory requirements and security standards, while continually delivering innovative offerings that connect with customers. Customers' diaital journeys with their bank cannot hit any speed bumps.

This is where quality engineering (QE) becomes critical. In the next few years GenAl will have a profound effect on organizations' QE functions – and vice versa. That's because GenAl is not only helping to move code through the system faster – it's accelerating a trend that was already starting to take shape. GenAl is helping to blend the roles of the developer and the quality engineer. Software processes have progressed a long way from the days when developer and quality engineer roles were completely separate and siloed. Testing no longer begins at the end of the software development lifecycle (SDLC), which historically slowed time to market by 6-8 weeks.

Now, the developer and quality engineer roles are changing rapidly, in large part due to the influence of GenAl. For example, having bots write more code shifts the demands on both developers and quality engineers. As coding itself becomes commoditized, developers' responsibilities shift to higherlevel, more strategic tasks such as architectural design, improving UX, incorporation of emerging technologies like blockchain, AR/VR, and edge computing. This in turn only increases the challenges quality engineering (QE) faces.

Quality engineering considerations are also being brought into more strategic discussions of business and IT. Not only will quality engineers play the crucial role of testing notoriously unreliable AI code – they'll be taking a seat at the table in the planning of software delivery strategies. GenAI's importance to the overall corporate mission elevates quality engineers' influence on critical topics ranging from IT transformation to CI/CD, to AI and ML model validation, to GenAI strategy and policy itself.

Ultimately, GenAl is changing the game for QE – putting more responsibility on the quality engineers' shoulders and more pressure on organizations to adapt.

SO, THE QUESTIONS ARE...

Is your organization ready for the necessary QE transformation being driven by GenAI?

Is your QE organization helping you get the most out of your corporate GenAl strategy?

Are you making the necessary investments to ensure that your GenAl strategy succeeds?

If the answer to any of the above is no, then it's time to keep reading.

GENERATIVE AI's IMPACT ON QE



GenAl is starting to play a more significant role in QE practices. According to IDC's latest forecast, the

Automated Software Quality (ASQ) market is set to grow at a CAGR of 12.7% through 2028*.

It indicates this growth reflects two interrelated trends: as 40% of GenAl-generated code requires remediation, the ASQ market must evolve to both test GenAldeveloped code and incorporate GenAl capabilities into testing tools themselves. This dual transformation is driving significant investment needs in advanced QE capabilities.

Our survey of 200 quality engineering professionals at financial service providers reveals their expectations and planned implementations. The findings show that QE professionals anticipate immediate impact on key technical KPIs, including improved automated test generations, data augmentation, continuous learning, efficiency in test maintenance and reduced human bias.



Top 5 Anticipated Benefits from Integrating GenAl into the QE Process

Source: Coleman Parks "Quality Times" survey done on behalf of Amdocs in 2024, respondents include 200 Global ClOs, Heads or Directors of Quality Engineering, and SDETs working in Financial Services institutions.

GENERATIVE AI's IMPACT ON QE



ven more importantly, QE professionals believe GenAl will impact KPIs the C-suite cares about – like higher customer satisfaction ratings, faster product development cycles, and increases in the number of quality issues identified before product release.

Top 5 Anticipated Measurable Outcomes When Integrating GenAl into Quality Engineering Processes



Source: Coleman Parks "Quality Times" survey done on behalf of Amdocs in 2024, respondents include 200 Global CIOs, Heads or Directors of Quality Engineering, and SDETs working in Financial Services institutions.

Yet while these anticipated benefits make a compelling case for GenAl adoption in QE, success requires careful attention to implementation.

GENERATIVE AI

QE'S IMPACT ON GENERATIVE AI



espite GenAl's promise, it won't succeed without significant attention from the QE side.

Here's the problem: Al-generated code is often inaccurate. In the Stack Overflow 2024 Developer Survey, 76% of respondents say they already are or are planning to use Al tools in their development process, but only 43% said they trust the accuracy of Al outputs.

Without intensive testing, mistakes in Al-generated code compound as lines of code flow through the software delivery system. There's inherent risk of trusting Al to generate code without quality oversight throughout the SDLC. The takeaway here is that if GenAl fails at this level, it exposes organizations to revenue losses, lost or corrupted data, security breaches, legal exposures, and hits to their brand reputations.

QE IN FINANCIAL SERVICES

Here are some GenAl use case examples in the financial services field:

REGULATORY-COMPLIANT SYNTHETIC DATA GENERATION

GenAl generates diverse, anonymized datasets that mimic real-world financial transactions for testing systems without breaching privacy regulations.

REGTECH QA AGENTS

Al agents interpret evolving regulatory requirements (e.g., GDPR, Basel III, DORA, PSD3), generate a corresponding test plan and test cases, and validate compliance automatically.

EDGE CASE SCENARIO SIMULATION

GenAl creates rare but critical test cases (e.g., financial crises, market anomalies) to stress-test systems, ensuring they can handle extreme load conditions. Additionally, GenAl generates chaos engineering scenarios, such as cyber-attacks and unexpected system failures, to identify and address potential weaknesses in system resilience.



ADOPT A MATURITY MODEL

The integration of GenAl into quality engineering infrastructures is set to transform testing methodologies and elevate the capabilities of testers. This vision outlines three maturity levels for leveraging GenAl, each representing a step forward in the sophistication and autonomy of testing processes.

BASELINE MATURITY

Assisted GenAl QE: GenAl serves as a knowledge base and instructional guide for testers, using simple machine learning models to answer questions and provide training. Human-driven interactions focus on low-criticality tasks, knowledge sharing, and basic testing support.

MID MATURITY

Augmented GenAl QE: GenAl actively supports standard testing workflows, becoming a collaborative partner with testers to enable semiautonomous processes. It enhances procedures by automating elements of test planning, test design, and workflow management.

HIGH MATURITY

Agentic AI QE: GenAl agents operate autonomously, handling complex testing tasks, optimizing environments, and managing configurations and data. By taking on critical quality engineering tasks, such as testing accelerators and orchestrating multi-layered environments, GenAl allows human testers to focus on strategy and oversight.



GenAl-Driven Quality Engineering Use Cases

Assisted GenAl QE

1. Knowledge Retrieval: Utilizing NLP natural language processing (NLP).

2. Interactive Learning: For hands-on learning and real-time support.

3. Automated Reporting: Summarizes test results and generates insightful summaries.

Augmented GenAl QE

1. Automated Test Design and Planning:

Creates test plans and designs test cases based on specifications and historical data.

2. Dynamic Workflow Optimization: Analyzes workflows and suggests optimizations.

3. Intelligent Test Execution: Autonomously executes preconfigured tests, adapts to changes.

Agentic GenAl QE

1. Autonomous Testing Agents: GenAl agents manage full test cycles, eg. validation, regression testing, bug tracking.

2. Data Management and Environmental Configuration: Autonomously manages test data and configures environments.

3. Self-Optimizing Testing Pipelines: Uses reinforcement learning to optimize testing pipelines to improve efficiency and accuracy.

CREATE NEW QE ROLES

In the GenAl-driven Quality Engineering era, as organizations progress along the maturity model, test engineers will transition from **executors** to **cognitive orchestrators** of intelligent, autonomous systems. Their roles will focus on designing adaptive and smart prompts and ensuring Al tools align with business goals. This transformation becomes clear when we look at how the role has progressed over time.

For example, see how a quality engineer's skills have evolved to where they are today and where they are heading in the GenAl era:

| Era | Waterfall | Agile | DevOps | GenAl Assisted | GenAl Augmented | Agentic Al |
|-----------|--|--|---|--|---|---|
| | J. | ×↑ o× | ۲۰۰۶ ۲۰/۶ ۲۰/۶ | | | |
| Job Title | Manual Test Engineer | Test Automation Specialist | Full Stack Quality Engineer | Al Enhanced Quality Engineer | Cognitive Orchestrate | or Quality Engineer |
| Skills | Test case design, manual testing, defect logging, requirement validation | Scripted test automation, test data preparation | API testing, continuous integration, test-driven development | Advanced prompt engineering, works with NLP tools and GenAl- driven knowledge management systems | ML training, test automation, workflow analysis | Reinforcement learning, orchestration platforms and data management for system optimization and troubleshooting |
| Outcome | Functionally verification | Partial automation reduces manual effort but cannot handle rapid changes effectively | Full Automation regression coverage, improving speed and consistency, ensuring seamless releases | Leveraging GenAl to retrieve knowledge, perform basic testing tasks, and support team knowledge sharing | Leveraging GenAl in test design, planning, workflow management, and automatic test case generation | Managing and overseeing autonomous GenAl testing agents that execute E2E processes |

ELEVATE THE IMPORTANCE OF PROMPT ENGINEERING

As we rapidly enter the GenAl era, Prompt Engineering has become a critical skill for quality engineers. With Al systems evolving into co-creators and autonomous agents within QE workflows, this skill of crafting prompts has advanced to creating sophisticated, business-oriented inputs. In the financial industry, where compliance, security, and customer trust are paramount, prompting skills must address all these considerations.

Here's how this evolution is reshaping key aspects of QE in financial services:

HANDLING FINANCIAL COMPLEXITY WITH ACCURATE TEST COVERAGE

In financial services, test cases must address the intricacies of regulatory requirements, complexities high-value transactions, and the challenges of fraud detection. Al has made it possible to dynamically create test cases that reflect these challenges. **Today:** Prompts are used to create basic test scenarios for tasks like payment processing or user authentication.

In the future: Al-powered test cases will simulate nuanced financial scenarios, including complex transaction flows, compliance stress testing, and edge cases involving fraudulent activity.

Example: Generate 10 test scenarios for a cross-border payments system focusing on scenarios like anti-money laundering (AML) compliance failures, simultaneous high-value transactions, and system overload during high trading volume days.

This evolution ensures financial systems are resilient against vulnerabilities while meeting stringent regulatory requirements.

MANAGING DYNAMIC FINANCIAL SYSTEMS THROUGH AUTONOMOUS TESTING AGENTS

Financial systems often experience real-time changes due to market fluctuations, regulatory updates, or user demands. Autonomous testing agents enabled by prompt engineering can adapt to these changes instantly. **In the future:** Testing agents in financial services will use adaptive prompts to modify test cases based on regulatory changes, customer behavior patterns, or even intraday market activity.

Example: Set up a regression suite for the latest compliance update. Prioritize modules with recent code changes, high-risk trading algorithms flagged by operational risk teams, and user-reported issues in mobile banking applications. Include high-frequency trading scenarios for stress testing.

This approach ensures that systems remain secure and compliant, even as the financial landscape evolves rapidly.

ADDRESSING BIAS AND ROBUSTNESS THROUGH DATASETS AND MACHINE LEARNING MODEL TESTING

Financial institutions rely on Al models for tasks like fraud detection, credit risk scoring, and investment analysis. These models must be tested for robustness, fairness, and accuracy using realistic and compliant datasets. Fairness includes testing for bias by evaluating models to ensure they do not unfairly disadvantage any group. **In the future:** Prompt engineering will enable quality engineers to create synthetic datasets that mimic real-world financial transactions while adhering to privacy regulations. Prompts will also test AI models for robustness against adversarial inputs, ensuring reliability and fairness.

Example: Generate a synthetic dataset of 1,000 banking transactions, including 10% of data representing anomalies to mimic fraudulent transaction patterns like unusually high deposits in dormant accounts.

This process helps financial institutions validate AI models while ensuring they perform accurately across diverse and complex scenarios.

WHY PROMPT ENGINEERING IS ESSENTIAL IN FINANCIAL SERVICES

In financial services, even a minor error in an AI system can lead to significant financial losses, regulatory penalties, or reputational damage. Prompt engineering helps quality engineers mitigate these risks by:

- Detecting vulnerabilities in high-stakes systems like trading platforms and payment gateways.
- Ensuring AI models comply with regulatory standards, including those related to data privacy and anti-discrimination.
- Enhancing customer trust through reliable, secure, and transparent financial systems.

In the age of AI, quality isn't just a function; it's a strategic advantage. Prompt engineering is the bridge that connects advanced technology with the reliability and trust businesses need to thrive.

5

HONGE MAXIMIZS BUSINESS BENEFITS FROM GENERATIVE AI



HOW QE MAXIMIZES BUSINESS BENEFITS FROM GENERATIVE AI

LLM OPTIMIZATION AS A SERVICE

Beyond these prompt engineering capabilities, QE is set to play a transformative role in optimizing large language models (LLMs) for enhanced customer experiences. This is particularly critical for financial institutions, where AI applications in customer service, wealth management and fraud detection demand unprecedented accuracy and reliability. We describe this comprehensive approach as "LLM Optimization as a Service" (LLMOaaS).

LLMOaaS is set to emerge as the next wave in quality engineering, enabling banks to achieve 99.9% accuracy in AI responses while maintaining compliance with financial regulations. The approach ensures GenAI models deliver not only accurate answers but also relevant, context-aware and trustworthy interactions – all of which are critical for sensitive financial operations.

LLMOaaS will serve as a hybrid model combining both service and software elements. The service will enable organizations to pull any content, context and responses from GenAI applications, measure them against multifactor KPIs and feed these insights back into the optimization loop of the LLM. Quality engineers will be key to this transformation, blending advanced techniques, robust infrastructure and next-gen KPIs to ensure enterprise LLMs are continuously evaluated, tuned and optimized for peak performance. It will require them to transition from traditional software testing into roles that involve tuning, enhancing and validating the performance of LLMs across a variety of contexts.

By offering ongoing analysis and finetuning capabilities, LLMOaaS can help deliver consistently high-quality interactions and enhance the overall customer experience. Operating as a blend of human-driven quality engineering expertise and automated Al-driven evaluations, it will continually push the boundaries of what LLMs can achieve.

The primary methods for LLM optimization will revolve around context understanding, content evaluation and continuous feedback loops, combining traditional quality engineering practices with AI-specific metrics and tuning strategies. Below are the key components of the approach for LLM optimization:

1. Contextual Relevance Evaluation: Testing LLMs for how well they adapt responses based on factors like user intent, past interactions, and the specific needs of the conversation.

2. Content Quality and Integrity Assurance: Evaluating content accuracy, appropriateness, and tone. Fact-checking integrations, sentiment analysis, and entity recognition pipelines will serve as automated methods to evaluate the truthfulness, sentiment, and logical coherence of LLM outputs.

3. Adaptive Learning and Fine-Tuning: By leveraging reinforcement learning from human feedback (RLHF), the model can be optimized for higher performance in areas where it historically struggled, creating an agile quality engineering ecosystem where models improve continually.

4. Bias and Ethical Quality Assurance (QA) Testing: As LLMs interact across diverse contexts and users, ensuring they provide equitable, unbiased answers is essential. This will involve creating robust bias detection and mitigation frameworks that can be tuned based on specific business or ethical standards.

As QE professionals and organizations transition to this new vision, LLMOaaS will empower them to design GenAI applications that not only answer questions but do so in ways that improve the customer experience.

HOW QE MAXIMIZES BUSINESS BENEFITS FROM GENERATIVE AI

ESTABLISHING RESPONSIBLE AND ETHICAL AI PRACTICES

As GenAl capabilities expand across all aspects of QE, establishing responsible and ethical AI practices becomes essential. In a GenAlenhanced future, quality engineers will transform into being knowledge facilitators, automation designers, and AI supervisors, working closely with intelligent systems to ensure the quality of next-gen applications.

Yet for this transformation to succeed, establishing AI policies becomes essential. This is where quality engineers will play a key role, ensuring that AI models operate responsibly, ethically, and in compliance with both organizational and regulatory standards.

Importantly, this will require a blend of technical expertise, governance, oversight, and proactive intervention to maintain trust in AI systems, which will be reflected in the following areas:

Ethical AI Governance

- Policy implementation Translating organizational AI ethics policies into actionable quality assurance (QA) frameworks or validations.
- Fairness audits Validating that AI decisions do not disproportionately impact specific groups or violate fairness principles.
- Explainability testing Ensuring AI models provide interpretable and transparent outputs to meet regulatory and user demands.

Regulatory compliance – Compliance validation and traceability assurance. When the use of Agentic AI expands, organizations will need to ensure all AI decisions are logged and traceable back to the intent.

Policy enforcement automation – Similar to DevOps where pipeline automation ensures continuous integration, GenAI will have to use AI model gatekeepers which will create quality gates that evaluate models before deployment to ensure adherence to all items. While these governance practices apply to current LLM implementations, they become even more critical as Agentic AI evolves to act as an "inorganic colleague." At this advanced stage, quality engineers will need to supervise AI systems in the following ways to maintain appropriate checks and balances:

- Police AI behaviors in real time.
- Enforce policies dynamically based on evolving standards.
- Provide automated, explainable audits for regulators and stakeholders.

HOW QE MAXIMIZES BUSINESS BENEFITS FROM GENERATIVE AI



How can all this work in practice? Imagine a scenario where GenAI takes the traditional role of QE to another level, enabling a bank to provide a completely different, knowledgeable, nuanced experience from an agentic agent that helps to manage inquiries and select banking activities for retail and business customers. Imagine how the bank could:

Introduce a customer to a trusted Agentic agent who matches the company brand persona, but is tailored to the specific customer through LLM Optimization as a Service. Provide an end-of-year summary for customers on the state of their finances through this agentic agent, highlighting areas of interest to help plan one's financial future, taking into account relationships with family dependents, obligations, risk tolerance and preferences. Leverage AI policing policies to ensure privacy and preservation of sensitive data, meeting regulatory requirements and giving their customers peace of mind.



CONCLUSION



s financial organizations embrace the transformative

power of GenAI, the role of QE must evolve to meet the new complexities it brings. GenAI is not just another tool to be bolted on and forgotten – it's a catalyst for rethinking how testing, optimization, and assurance are approached. While the opportunities for innovation appear limitless, its impact on testing environments and the entire modern software ecosystem is anything but simple.

To scale and accelerate these transformation efforts effectively, organizations need to make sure quality engineering discussions are taking place at the highest level. One way to do this is to create a new role overseeing GenAl's integration into QE. This would ensure operational excellence while also aligning innovation with broader business objectives. This leadership role could become the architect of trust, ensuring systems are reliable, secure, and prepared for the demands of tomorrow. For financial organizations, where the stakes are highest, organizations should consider establishing a Chief Quality Officer who can unlock new business opportunities by leveraging and continuously optimizing GenAl to its fullest potential. In an era of accelerated change, this move would provide peace of mind and a competitive edge.

The time to act is now. By investing in Quality Engineering, appointing visionary leaders, and embracing GenAl's potential, organizations can turn challenges into opportunities. The future belongs to those who prioritize quality – not just as a function, but as a guiding principle for innovation and growth.

The Amdocs Quality Engineering brAln framework increases the level of autonomy and trust in your AI models, which enables your QE organization to go from Assisted to Agentic AI. The adoption of the brAln in our practices has resulted in our customers recognizing a **33% decrease** in Testing Design time, **25% decrease** in Automation Development time (MTTC), **50% improvement** in Test Coverage Optimization, **60% decrease** in Testing Certification Time, and **65% decrease** in Tester Onboarding Time.

<u>Talk to the Amdocs team</u> to discover how we can support your QE transformation and help you stay ahead in the rapidly evolving tech landscape!

Amdocs helps those who build the future to make it amazing. With our market-leading portfolio of software products and services, we unlock our customers' innovative potential, empowering them to provide next-generation experiences for both the individual end user and enterprise customers. Our employees around the globe are here to accelerate financial institutions' migration to the cloud, enable them to differentiate in the digital era, and automate their operations.

Listed on the NASDAQ Global Select Market, Amdocs had revenue of \$5.00 billion in fiscal 2024.

For more information, visit Amdocs at www.amdocs.com

amdocs STUDIOS

Experience & Digital Engineering | Data & GenAl | Cloud | Quality Engineering

