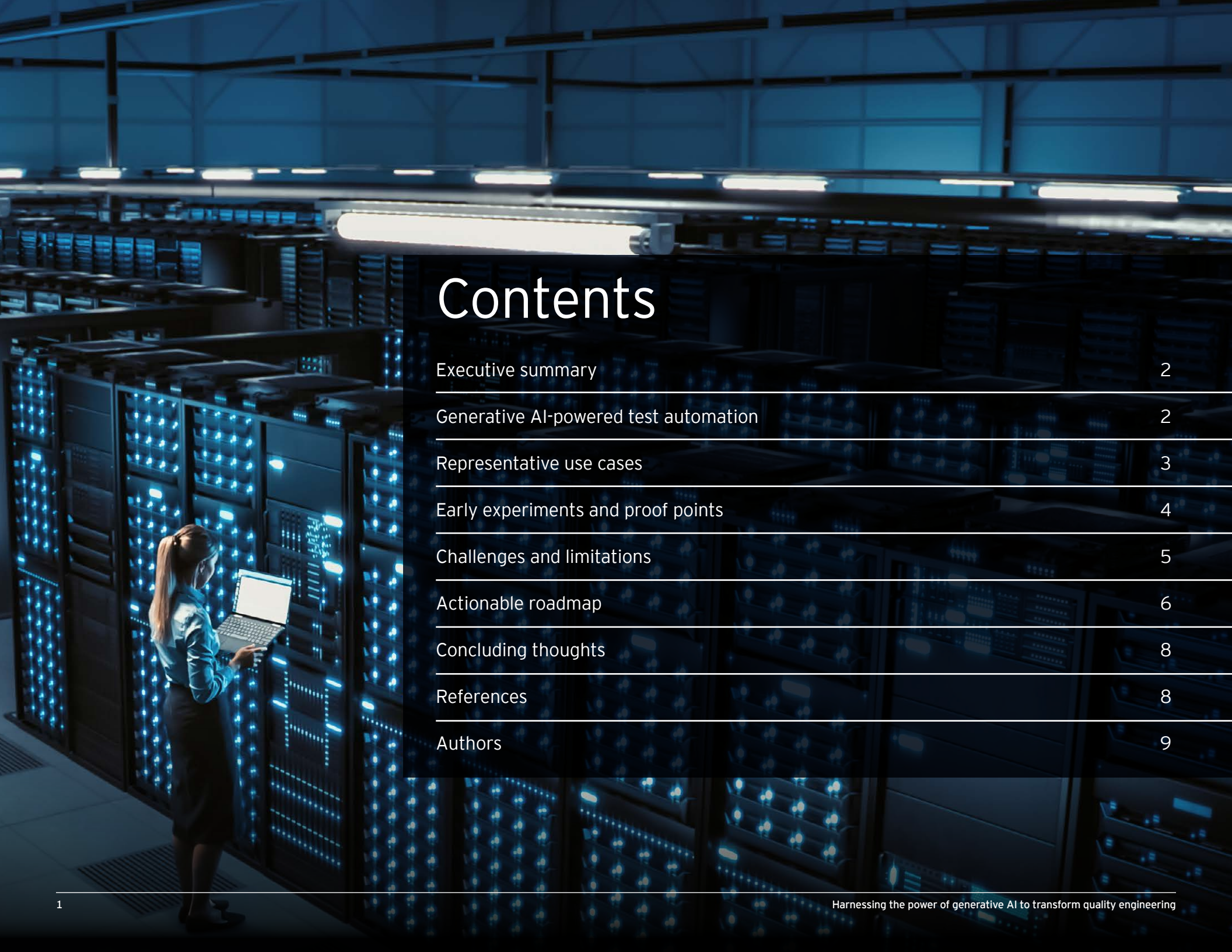
Harnessing the power of generative AI to transform quality engineering

EY
Building a better working world



Contents

Executive summary

Generative AI, commonly referred to as GenAI, surpasses traditional artificial intelligence (AI) capabilities by empowering machines to not only comprehend information but also create, imagine and innovate. The latest EY Reimagining Industry Futures Study[1] confirms GenAI's status as a breakthrough technology, with 43% of the 1,405 enterprises surveyed investing in it. This represents merely the tip of the iceberg, with boundless possibilities awaiting exploration. Specifically, GenAI technology brings immense value to software test lifecycle management in delivering high-quality software products, expediting test cycles, enhancing test coverage and boosting tester productivity.

This paper explores the prospective future of software quality test automation, leveraging the transformative capabilities of GenAI technology. Based on early experiments, it outlines the priorities of representative use cases. The paper discusses the challenges and limitations and puts forward actionable recommendations for enterprises, outlining immediate, midterm and long-term strategies to seamlessly integrate GenAI capabilities into their testing practices.

GenAI-powered test automation

Historically, AI has served to complement human efforts by automating repetitive tasks, with AI-driven regression suite automation serving as an example. Earlier AI-based defect classification systems paved the way for the development of more advanced approaches using machine learning (ML) for defect prediction and analysis in contemporary software testing methods. The recent evolution in GenAI technology autonomously generates test plans and test designs, and summarizes the test execution insights, resulting in substantial improvements in productivity.

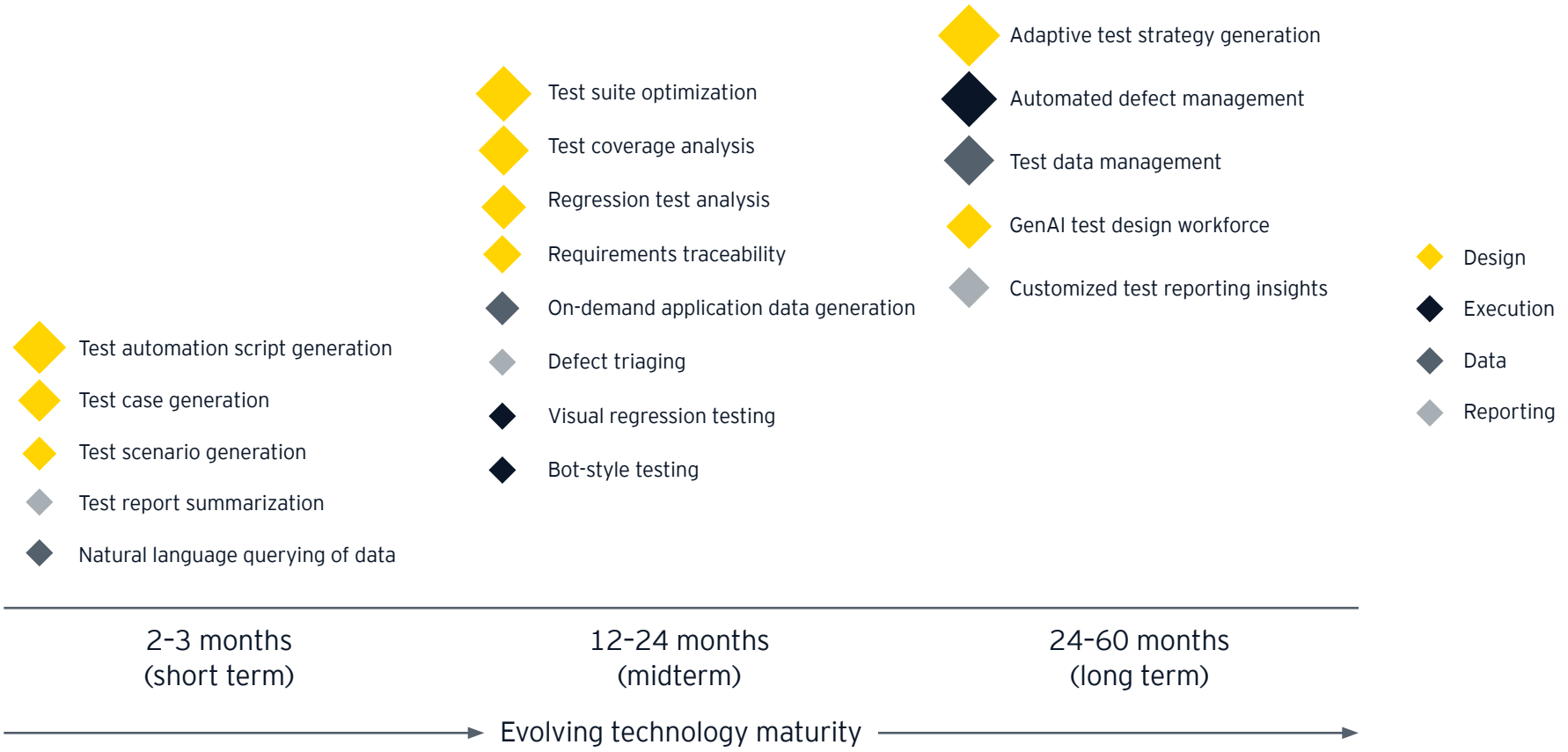
Not only do these systems drive efficiency, but they also possess self-healing abilities. They can independently identify and rectify issues that arise during test execution, thereby enhancing the overall quality and resilience of the product or service.

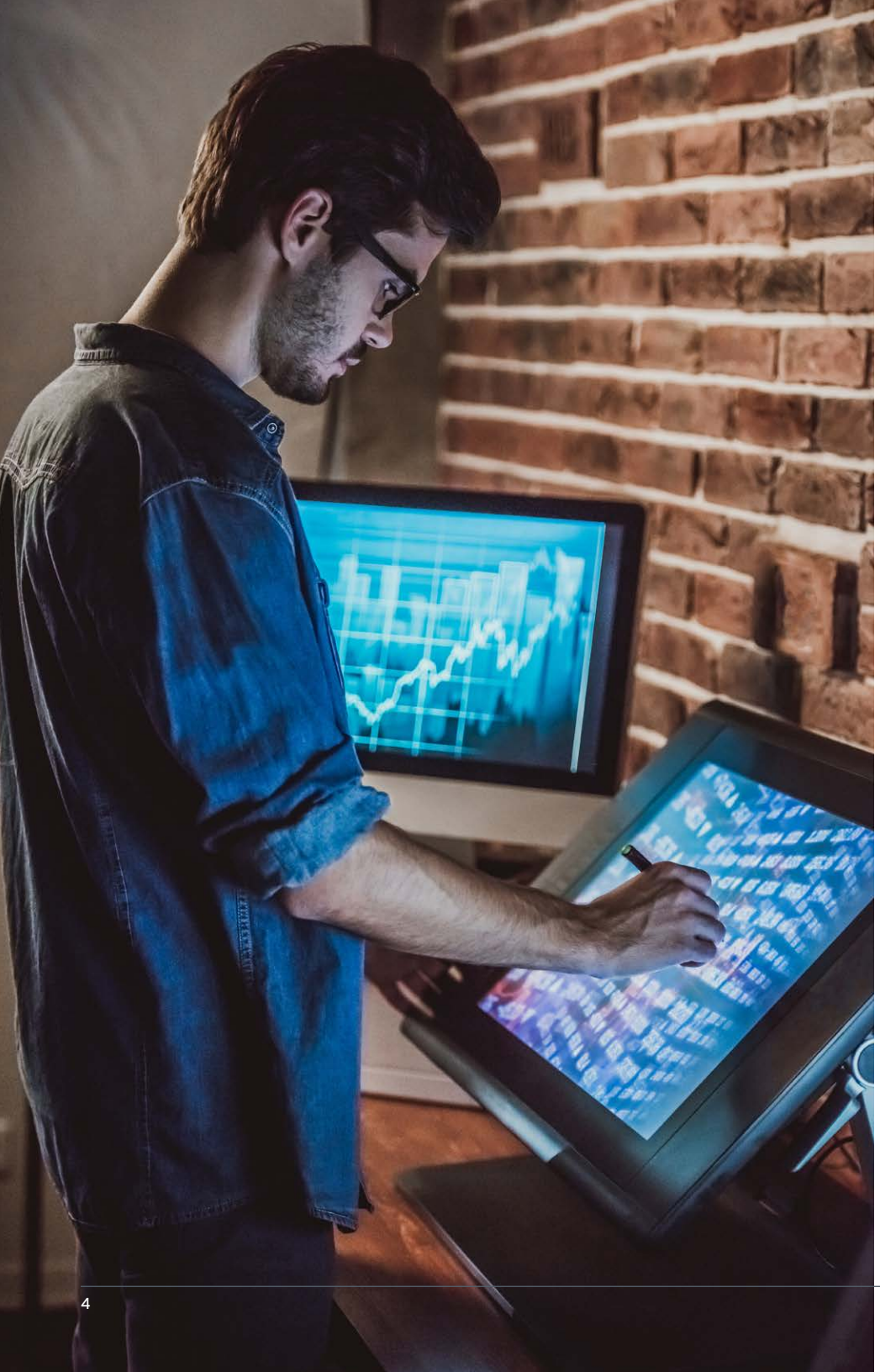
Representative use cases

The figure below illustrates 18 representative use cases of GenAI technology in a maturity continuum across the software testing lifecycle that can be implemented in the short, mid and long term. As the technology matures over time, it is anticipated to reduce the need for human intervention and would have the ability to seamlessly integrate with the enterprise automation platforms and tools. Value potential, represented by the size of the diamond, denotes the feature or capability impact on the final quality of the product or service.

Representative GenAI testing use cases in a maturity continuum

Bigger diamonds represent higher value potential.





Early experiments and proof points

Following are some of the proof points from our early experimentations for global clients spanning varied sectors.

Proof point 1	**US state government agency**	▸ Generated and optimized UAT test cases based on existing SIT test cases ▸ Approximately 50% reduction in test cases from what was originally planned
Proof point 2	**Global retirement and life insurance organization**	▸ Generated test cases from requirements/user stories and built test automation scripts ▸ Approximately 80% accuracy achieved using domain knowledge and iterative prompt engineering
Proof point 3	**World's leading automobile firm**	▸ Legacy modernization application testing from COBOL to Java ▸ Generated test cases through analysis of COBOL code, existing test case repository, and requirements documentation ▸ Approximately 40% reduction in manual test design efforts

Challenges and limitations

The findings from the early experiments indicate that GenAI holds promise for automating certain testing tasks, improving test coverage, and enhancing the effectiveness and efficiency of the overall testing process. However, it is important to consider the following limitations and challenges associated with GenAI technology.

Not a panacea to address all problems

GenAI is not a one-size-fits-all solution for addressing all testing needs. While it can be suitable for certain testing requirements, there are situations where traditional testing approaches may be more appropriate and effective. Careful evaluation and a broad solution approach considering alternative testing methodologies and human involvement are necessary to provide holistic test lifecycle management.

Data privacy and explainability concerns

Lack of transparency on the usage of GenAI hampers the assessment and resolution of concerns regarding output inaccuracy, data privacy and user consent. In regulated industries like financial services or health care, compliance with ethical guidelines and legal regulations is critical. Organizations should document and report on GenAI-powered testing approaches, including risk mitigation measures and testing methodologies.

Addressing bias in generated output

The application of GenAI technology can lead to inherent bias, discrimination and promote unsafe decisions, which may originate from the training data itself. Organizations must proactively establish ethical frameworks and guidelines to mitigate risks and have measures in place to validate the generated test outputs. The recent executive order by President Joe Biden[2] underscores the critical need for the development of standards, tools and tests to ensure the safety, security and trustworthiness of AI systems. This noteworthy step reflects the growing acknowledgment of the potential implications and risks associated with AI technologies.

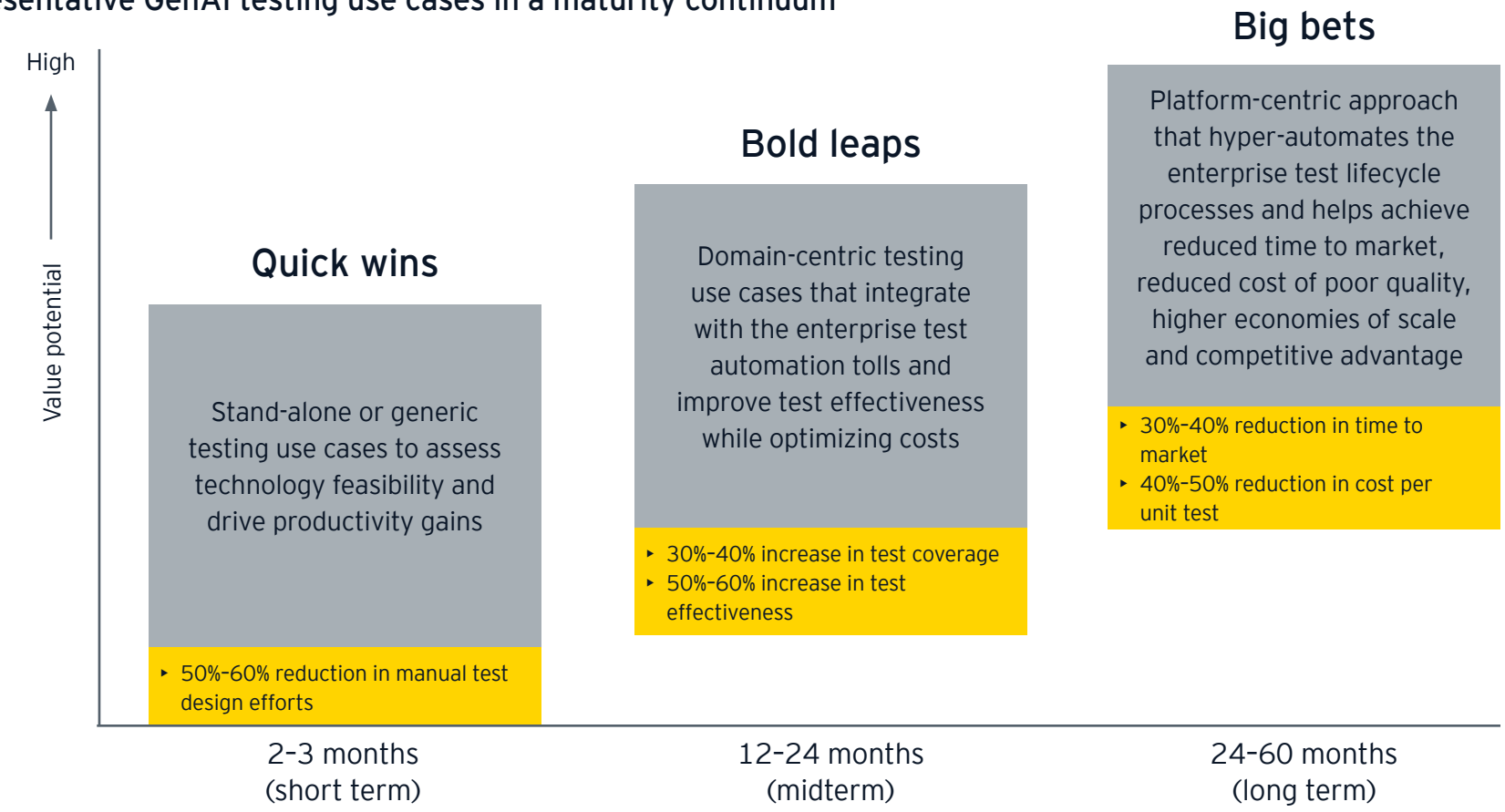
Need for human oversight and validation

Testing often involves human expertise, domain knowledge and judgment to identify relevant test cases, prioritize testing areas and assess the significance of certain scenarios. GenAI, while helpful, may still lack the human intuition, domain context and critical thinking required for effective test planning, design, execution and reporting. Having a human in the loop to validate the output accuracy becomes inevitable.

Actionable roadmap

Organizations must adopt a mindset of thinking big, starting small, embracing failure as a learning opportunity, and scaling their GenAI-enabled testing initiatives to foster continuous innovation and growth. The figure below highlights an indicative roadmap demonstrating the three-phase approach for GenAI adoption in software testing. This roadmap must be tailored based on the organizational goals, adoption readiness and impact potential to quality processes.

Representative GenAI testing use cases in a maturity continuum



This table summarizes illustrative figures across the three phrases. The key considerations are projections.

Key considerations	Quick wins	Bold leaps	Big bets
Engagement type	Proof of concept/pilot	Scale at a business unit or large engagement	Scale across the enterprise
Timeline	2–3 months	12–24 months	24–60 months
Use cases	Generic	Domain-centric	Enterprise platform-centric
Integration	Stand-alone	Integration with requirement management and test automation tools	Integration with enterprise processes, systems and tools
Data management	None	Domain-centric training data, vector databases, etc.	Enterprise-level training data, specialized data vector databases, knowledge graphs, etc.
Suggested model	Pretrained	Pretrained or fine-tuned	Pretrained, fine-tuned or custom-built
Model management	None	Required in provider-managed or self-managed infrastructure	Required in provider-managed or self-managed infrastructure
Talent management	Highly skilled developers and testers	Highly skilled software developers, functional domain tester SMEs, GenAI-trained automation testers, GenAI engineers, data scientists, DevOps engineers	Highly skilled software developers, enterprise domain tester SMEs, GenAI-trained automation testers, GenAI engineers, data scientists, LLMOps engineers
Change management	Upskill existing talent on GenAI tools usage	Upskill talent on GenAI-enabled, domain-centric solutions or transformed processes; comprehend AI vs. human roles in undertaking the testing tasks or job roles	Upskill talent on GenAI enabled enterprise platform-centric solutions or transformed processes; comprehend AI vs. human roles in undertaking the testing tasks or job roles
AI governance	Minimal oversight	Adequate oversight on responsible AI practices	Dedicated COE to provide oversight on responsible AI adoption frameworks and practices; integrate the responsible AI governance guardrails as part of the testing workflow tasks or job roles
Value projection	50%–60% reduction in manual test design efforts	▸ 30%–40% increase in test coverage ▸ 50%–60% increase in test effectiveness	▸ 30%–40% reduction in time to market ▸ 40%–50% reduction in cost per unit test

Concluding thoughts

According to the Gartner Hype Cycle 2023 for AI,[3] the GenAI movement encompasses two pivotal aspects driving the development of more powerful AI systems. On one side, there are the innovations that will be fueled by GenAI capabilities, leading to advancements across various domains including software testing. On the other side, there are the innovations that will fuel the progression and refinement of GenAI itself, enabling it to become even more capable and impactful. Apparently, the future of GenAI for testing holds tremendous potential for advancements in testing effectiveness, efficiency and coverage. Enhanced realism, adaptive generation, domain-specific models, collaboration with human testers, self-learning and self-healing through continuous feedback loop, and seamless integration with enterprise requirement and/or test management tools will shape the next generation of GenAI for testing. To stay relevant, organizations are advised to initiate quick win strategies and gradually increase investments in line with the evolving business needs.

References

[1] EY Reimagining Industry Futures Study 2024, assets.ey.com/content/dam/ey-sites/ey-com/en_gl/topics/tmt/ey-reimagining-industry-futures-study-2024-report.pdf, EYGM Limited, 2024.

[2] "FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence," whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/, The White House, 30 October 2023.

[3] "What's New in Artificial Intelligence from the 2023 Gartner Hype Cycle," gartner.com/en/articles/what-s-new-in-artificial-intelligence-from-the-2023-gartner-hype-cycle, *Gartner website*, 17 August 2023.

Ernst & Young LLP and EY Global Delivery Services authors

Narayan Arunan Surya
Managing Director, Technology Consulting,
EY Global Delivery Services

Santhi Duraisamy
Managing Director, Technology Consulting,
EY Global Delivery Services

Akila Narayanan
Senior Manager, Other Consulting,
EY Global Delivery Services

Dhanesh Kumar
Senior Manager, Technology Consulting,
EY Global Delivery Services

Justin Hanke
Senior Manager, Technology Consulting,
Ernst & Young LLP

EY | Building a better working world

EY exists to build a better working world, helping to create long-term value for clients, people and society and build trust in the capital markets.

Enabled by data and technology, diverse EY teams in over 150 countries provide trust through assurance and help clients grow, transform and operate.

Working across assurance, consulting, law, strategy, tax and transactions, EY teams ask better questions to find new answers for the complex issues facing our world today.

US SCORE no. 24649-241US_2
BSC no. 2407-84079-CS
ED None.

ey.com