Cognizant Digital Systems & Technology

How QA Ensures that Enterprise AI Initiatives Succeed

The euphoria around artificial intelligence (AI) focuses primarily on what it can do, leaving the hard work for expert teams to sort through. A curated quality assurance (QA) strategy, focused on parameters such as data, algorithm, biases and digital ethics can ensure that AI initiatives deliver.

Executive Summary

As with most breakthrough technologies, AI was initially received with heightened euphoria, followed by frenzied adoption. After nearly seven decades, AI now finds itself nearly commonplace, as self-driving cars, intelligent toothbrushes and personal digital assistants are infused into everyday life. By analyzing data in real time, AI systems can now make autonomous decisions that often bear significantly on humans, such as estimating creditworthiness for a loan, or shortlisting résumés based on a job description.

While the focus has been largely around what AI can do, assurance for AI has been mostly unknown. Success of an AI application depends on its ability to analyze information and then make a judgement call to either do nothing, act autonomously, or raise a flag for human intervention. Moreover, AI applications must continuously learn from past incidents/data to hone future decisions. As AI applications often make decisions regarding humans, it is important to keep biases at bay. (Read our report "Making AI Responsible & Effective.")

January 2020

In the rush to join the bandwagon, enterprises tend to overlook the two building blocks that differentiate an AI application from any other software – data and algorithm. Instead, they treat AI applications just as they would traditional software, ending up with encoded biases, false negatives or positives and, in extreme cases, an AI program that goes rogue.[1]

Clearly, a few fundamentally unique parameters determine the quality of the AI application, for which a traditional approach will not work.

This white paper discusses how AI applications are different from traditional software and how end-to-end quality assurance (QA) can help enterprises ensure that their AI initiatives succeed.

AI applications are different

AI applications are fundamentally different from traditional software (see Figure 1) in that they have:

- **No definite input.** Contrary to traditional software, where definite statements are inputs, AI applications work with a range of probabilities. For instance, traditional software tracking physical activity would require "if statements" about walking, jogging or running. On the other hand, an AI application would need varying ranges of speed labeled walking, jogging or running. Now, if a new activity, such as cycling, needs to be tracked, the traditional software would require another clear-cut "if statement," whereas a different speed range would be the input for the AI application.
- **No definite output.** As a continuous self-learning system, AI applications tend to evolve better with time, hence defeating the notion of "expected outcomes" that govern traditional software. For instance, over time, the AI system would learn and narrow down the speed ranges to more specific activities based on user profiles (i.e., walking speeds based on user age). It could then use these speeds to predict the optimum level of activity as per age. The output here is a predictive range rather than a definite value.
- **Iterative lifecycle.** Due to non-deterministic input and output, the AI lifecycle is iterative, which means the data that it generates as output becomes input for future instances. Hence, the algorithm works in a continuous loop and learns from historical instances to enhance quality of predictions. This is in stark contrast to the traditional software lifecycle, which is sequential.
- **Propensity for bias.** If data such as speed and user age are being fed to the AI program, it could learn and identify patterns over time that could make it assume that elderly people walk slower than younger ones. It could even develop a propensity for bias, where it classifies a young person as elderly if he/she walks at a slower pace.

Traditional vs. AI software development

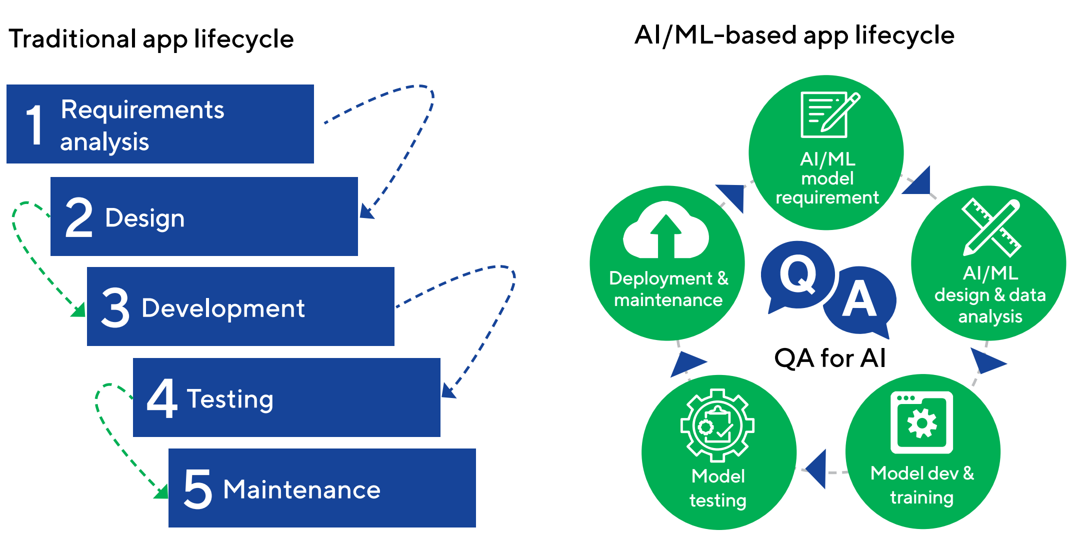


Figure 1

To select the right data sets, QA teams should check if the features add value to the expected output. For instance, if race and sex are being considered for creditworthiness, then biases in the data should be eliminated. QA should also weed out outliers—say data regarding number of siblings for creditworthiness that might contribute to noise and impact quality of outcome.

Pivots for assuring AI quality

- Even though immensely promising, AI programs must be viable from a business standpoint. In other words, AI for the sake of AI will not work. For instance, in 2012 Netflix paid $1 million for an algorithm that could have improved recommendations by 10%. However, the algorithm was shelved due to the mammoth engineering effort it would have needed to be actually implemented on Netflix's platform.[2]

Once enterprises identify the use case for AI, they need a QA strategy that focuses on:

- **Data.** It is crucial to identify correct data sets for training and testing. This will help arrest bias that might influence outcome. For instance, if an AI program calculating creditworthiness works with data factors such as income, age, spending capacity, sex and race, then the algorithm might become biased. (Learn more.)

  To select the right data sets, QA teams should check if the features add value to the expected output. For instance, if race and sex are being considered for creditworthiness, then biases in the data should be eliminated. QA should also weed out outliers; say data regarding number of siblings for creditworthiness that might contribute to noise and impact quality of outcome. Moreover, data should follow a certain trend for the AI to draw patterns and extract relevant information. (Read our white paper, "Business Assurance for Voice-Enabled Digital Assistants.")

- **Model.** This is the brain of the AI program; the model provides underlying business processes to ensure the AI works as intended. For instance, in AI-enabled testing software, the code is pushed to production only if it passes prerequisite validation tests. This will happen if the model makes an independent judgement call based on its knowledge of the tests' pass/fail criteria. (Learn more.)

  > QA teams should validate the model for:

  1. Sustainability: Identify the type of AI model (supervised/unsupervised), the technologies (Python, R and Spark) it is built on and a set of quality metrics aligned to business expectations, such as F1 score,[3] Confusion matrix[4] and ROC/AUC.[5]
  2. Feasibility: Regress the model out of development and qualify its production-readiness.
  3. Fairness: Conduct a deep dive analysis on the model to investigate the relationship between input and output. Consider sensitive attributes such as race, sex, socio-economic status and suggest corrective actions that can eliminate biases in data.

- **Process framework.** Following industry-benchmarked processes is important for ethical AI programs. For instance, in sanctioning a loan, the AI algorithm should ensure that the personal data provided by the applicant is secured against pilferage. The AI program should adhere to international guidelines such as General Data Protection Regulations (GDPR).

  To assure process frameworks, QA teams need to define the business value expected from AI and identify metrics and measurement techniques (using industry standards such as F1 score, Confusion matrix and ROC). The process should be assessed if it supports algorithms and data processing pipelines use related to business problems, or if it supports processing of structured and unstructured data. Data ingestion from unstructured and structured data sources can be handled with standard tools such as Talend, Logstash, etc. Moreover, QA teams should validate the AI model for governance, digital ethics (as in as in fairness of loans approved to all, irrespective of race or other factors) and robustness.

- **Performance and security.** Since AI programs work with classified data and compliance-mandated business processes, performance and security are crucial. For example, the model should be retrained within an allowable batch window, which should not take too long. The model should be robust enough to handle spikes or poor quality of data. Security should be validated to address vulnerability. (Read our white paper, "Applying Machine Learning to Boost Digital Business Performance.")

  QA teams should assess elements such as predictive response time for peak user loads or transactions and assess the model against standard performance tuning aspects such as CPU usage, memory consumption, etc.

Orchestrating AI quality

An AI program cannot rationalize its output, which means it is oblivious to internalized biases or deviation from industry benchmarks. This could undermine the intent and effort put into AI programs. QA needs to assume accountability of the output generated by AI programs by ensuring that the data, model, and process framework along with performance and security meet business expectations.

This requires an end-to-end approach to QA, through which issues such as biases are addressed as the AI algorithm is developed and trained. By embedding QA in every step of the lifecycle, AI applications are trained better and tested earlier for anomalies that may be difficult to rectify if discovered later. This typically requires a highly automated approach to QA, which feeds the AI multiple data sets pooled from

QA teams should assess elements such as predictive response time for peak user loads or transactions and assess the model against standard performance tuning aspects such as CPU usage, memory consumption, etc.

various resources, and limits human intervention to minimize human biases.

A platform-based approach works best, providing QA with a vantage point of an orchestrator, and enabling it to enforce quality standards at each stage in a console-like manner. (Learn more.)

Case in point

A leading telecommunications provider in the U.S. used a prediction algorithm for root-cause network analysis. The AI model was accurate in the lab, but when applied in the real world, its accuracy declined by 15%. Training data was a subset of 50 applications' data captured a few months back, while the current production application data was showing different data patterns compared to what was captured within training data sets. This was due to a misalignment between data science and production teams.

We proposed verifying and validating the prediction model during the building phase with data fitment analysis that identified the right attributes and ensured that the training data used in the lab reflected real-world production performance, thereby reducing accuracy variance.

The team also validated the prediction model with various sets of curated data (including features like defect description, severity, and priority) to identify biases and fairness. These features helped the client increase the accuracy of its prediction model to 74%.

The way forward

To deliver a seamless digital experience, AI applications must be built on a multi-layered architecture of technologies to accommodate data provided by third-party players.

With an ever-expanding ecosystem of stakeholders, QA teams need to evolve from guardian of quality to custodian, ensuring superior data quality across touchpoints, for all stakeholders in the ecosystem.

Quality orchestration is the way forward to assuring complex AI applications. Access to QA talent that is technologically sound and domain-oriented is critical to this evolution.

The agenda for enterprise QA organizations is to reskill and upskill resources to address the exponentially rising quality needs of complex digital technologies, such as AI.

With an ever-expanding ecosystem of stakeholders, QA teams need to evolve from guardian of quality to custodian, ensuring superior data quality across touchpoints, for all stakeholders in the ecosystem. Quality orchestration is the way forward to assuring complex AI applications.

About the authors

Vikul Gupta

Market Leader for Digital Assurance, Cognizant

Vikul Gupta is the Market Leader for Digital Assurance within Cognizant's Quality Engineering & Assurance Practice. He has 20 years of experience in strategy, delivery and solutioning, with an extensive background in data analytics, DevOps, Cloud, infrastructure automation, mobile and IoT. An astute techno-strategist, Vikul works with various business units to help chart Cognizant's enterprise digital roadmap and define a service delivery approach that is aligned with product strategies. He has rich experience across roles ranging from a product developer to solution strategist, and he is an industry-renowned thought leader. He is a graduate of the National Institute of Technology, Surat. Vikul can be reached at Vikul.Gupta@cognizant.com | www.linkedin.com/in/vikul/.

Saravanan Palanivelu

Solution Architect/Data Scientist, Cognizant

Saravanan Palanivelu is a Solution Architect/Data Scientist in Cognizant Quality Engineering & Assurance Practice. Passionate about creating data-driven products to address business challenges, he is leading the effort for quality intelligence and predictive analytics/machine learning within Cognizant's Quality Engineering & Assurance Practice; quality insight is a key initiative that he drives as part of his focus on digital assurance initiatives. Saravanan has 13-plus years of experience in testing, analytics, machine learning solution architecture, product management, technical product planning, product innovation and market research of enterprise software products. Moreover, he is responsible for creating new machine-learning analytics solutions to improve applications' quality. Saravanan graduated from National Engineering College, Manonmaniam Sundaranar University. He can be reached at Saravanan.Palanivelu@cognizant.com | www.linkedin.com/in/saravanan-palanivelu-74939740/.

Vasanthkumar Velayudham

Solution Architect/Technologist, Cognizant

Vasanthkumar Velayudham is a Solution Architect/Technologist, leading the effort for quality intelligence and predictive analytics/machine learning within Cognizant Quality Engineering & Assurance Practice. Part of his focus on digital assurance initiatives is quality insight. Vasanth has 12-plus years of experience in testing, analytics, technical solution architecture, experimental design, product management, technical product planning, product innovation and market research of enterprise software products. He is also responsible for creating new machine-learning analytics solutions to impove applications' quality. Vasanth is a graduate of SSN College of Engineering, Anna University. He can be reached at Vasanthkumar.Velayudham@cognizant.com | www.linkedin.com/in/vasanthkumar-velayudham-4a34ab25/.

Endnotes

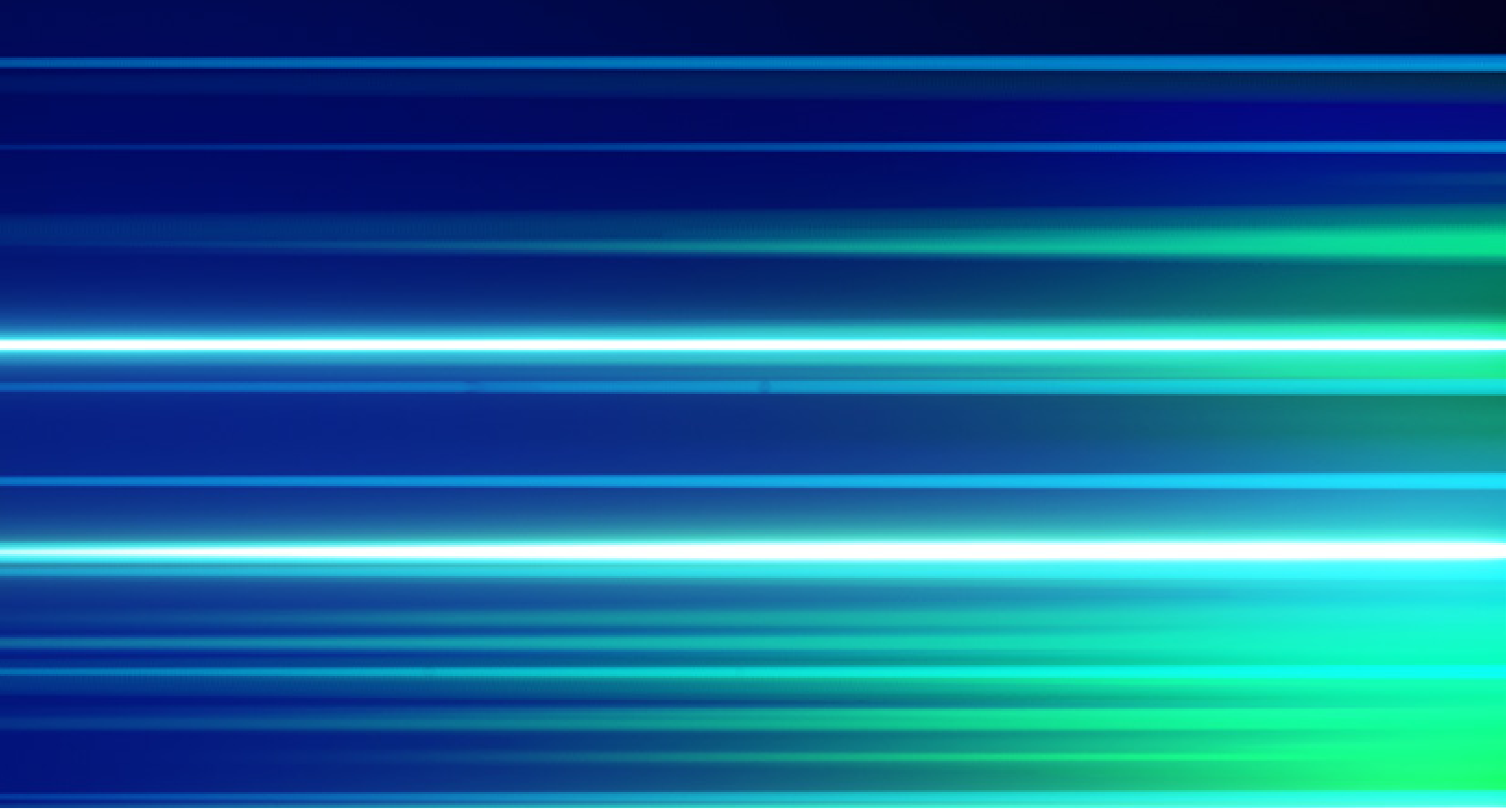
1 "Why Facebook Shut Down Its Artificial Intelligence Program That Went Rogue," Forbes.com, August 16, 2017, http://www.forbes.com/sites/quora/2017/08/16/why-facebook-shut-down-its-artificial-intelligence-program-that-went-rogue/#3ff9266a1710.

2 "Netflix Never Used Its $1 Million Algorithm Due To Engineering Costs," April 16, 2012, www.wired.com/2012/04/netflix-prize-costs.

3 https://en.wikipedia.org/wiki/F1_score.

4 https://en.wikipedia.org/wiki/Confusion_matrix.

5 https://en.wikipedia.org/wiki/Receiver_operating_characteristic.



About Cognizant Digital Systems & Technology

Cognizant Digital Systems & Technology works with clients to simplify, modernize and secure IT infrastructure and applications, unlocking the power trapped in their technology environments. We help clients create and evolve systems that meet the needs of the modern enterprise by delivering industry-leading standards of performance, cost savings and flexibility. To learn more, contact us at simplify@cognizant.com. You can also visit us at www.cognizant.com/cognizant-digital-systems-technology, or e-mail us at Inquiry@ cognizant.com.

About Cognizant QE&A

Cognizant Quality Engineering & Assurance (QE&A) helps businesses succeed in digital with an industry-aligned digital assurance proposition. With 800-plus clients across industry verticals and a global footprint, Cognizant is a recognized market leader in Quality Assurance. Cognizant's deep business and technology expertise helps our clients drive quality at speed with Zero-Touch QA. Cognizant's QA Hub™ ecosystem accelerates innovation by bringing together partners and communities to get quality right the first time. Learn more at www.cognizant.com/cognizant-digital-systems-technology/enterprise-quality-engineering-assurance

About Cognizant

Cognizant (Nasdaq-100: CTSH) is one of the world's leading professional services companies, transforming clients' business, operating and technology models for the digital era. Our unique industry-based, consultative approach helps clients envision, build and run more innovative and efficient businesses. Headquartered in the U.S., Cognizant is ranked 193 on the Fortune 500 and is consistently listed among the most admired companies in the world. Learn how Cognizant helps clients lead with digital at www.cognizant.com or follow us @Cognizant.

Cognizant

World Headquarters
500 Frank W. Burr Blvd.
Teaneck, NJ 07666 USA
Phone: +1 201 801 0233
Fax: +1 201 801 0243
Toll Free: +1 888 937 3277

European Headquarters
1 Kingdom Street
Paddington Central
London W2 6BD England
Phone: +44 (0) 20 7297 7600
Fax: +44 (0) 20 7121 0102

India Operations Headquarters
#5/535 Old Mahabalipuram Road
Okkiyam Pettai, Thoraipakkam
Chennai, 600 096 India
Phone: +91 (0) 44 4209 6000
Fax: +91 (0) 44 4209 6060

APAC Headquarters
1 Changi Business Park Crescent,
Plaza 8@CBP # 07-04/05/06,
Tower A, Singapore 486025
Phone: + 65 6812 4051
Fax: + 65 6324 4051

Codex 5074