Generative AI: A Catalyst for Data-driven Growth

Basavaraj Darawan
Senior Director, Analytics, WNS Triange
In the fast-paced landscape of modern enterprises, the importance of data has grown exponentially over the last three decades. A recent global data and analytics survey by Forrester Consulting, commissioned by WNS, revealed that 82 percent of organizations with advanced maturity levels witnessed positive year-on-year growth despite the pandemic.

Organizations are surrounded by vast and complex data assets from numerous structured and unstructured sources, including the Internet of Things (IoT) and streaming data. With mergers and acquisitions and global expansion becoming commonplace, businesses generate a constant stream of valuable data. This data is a treasure trove of insights that drive strategic decisions, and its proper utilization is crucial to achieving growth while optimizing costs.

However, despite its potential, much of this data remains underutilized, fragmented and inaccessible to those who need it the most. Given the sheer volume of data organizations possess, embracing modern data transformation initiatives becomes imperative to effectively manage and leverage this valuable asset.

The key challenges organizations face in data transformation revolve around data acquisition, integration, classification, tagging, quality, security, automation and data set management. Despite technological advancements and the integration of Artificial Intelligence (AI), these aspects persist as some of the most time-consuming elements in the data analytics value chain. As CXOs strive to foster growth and reduce costs, overcoming these obstacles to unlock the full potential of their data assets becomes a strategic priority.
Leveraging Generative AI for Data Management

Generative AI (Gen AI), powered by sophisticated Large Language Models (LLMs), is emerging as a transformative force in data-driven decision-making. Recent McKinsey research found that 90 percent of commercial leaders expect to harness Gen AI solutions "often" in the foreseeable future. Gen AI presents realistic solutions to data transformation challenges by introducing AI-assisted tools and technologies. This intersection of Gen AI with data analytics is set to re-define how businesses extract value from their data, making strategic growth decisions more achievable than ever.

Several areas are poised to benefit significantly from Gen AI. Data quality, in particular, will witness numerous novel features addressing data anomalies and outliers. Data quality is pivotal for decision-making in virtually all aspects of analytics, as it directly affects the accuracy and reliability of reports and AI models. High-quality data empowers AI models to make better predictions and yield more reliable outcomes, fostering trust and confidence among users. However, configuring data quality rules requires specialized technical and domain expertise.

Gen AI will empower data professionals by tackling the significant hurdles posed by labor-intensive processes inherent in configuring data quality systems. It will bring in Gen AI-assisted features that streamline the analysis, configuration and optimization of various data programs and rules required to validate and rectify data sets. The time-consuming tasks of dataset profiling and manual identification of data issues will be orchestrated through Gen AI-enabled automation.

Gen AI algorithms will conduct data quality analysis or assessment, unearthing anomalies, outliers and complexities within the data sets while quantifying their impact on the business. The analysis outcomes will lead Gen AI models to propose a comprehensive set of business and technical rules to address the identified data quality enhancement challenges. These suggested rules will be presented in native English, allowing users to understand and authorize their implementation.

For instance, in the insurance sector, Gen AI will proficiently analyze and illuminate the intricate interplay between policies and claims. With the help of synthetic data, Gen AI will lay bare the concealed relationships between various entities, identifying any references to missing policies for processed claims and other crucial parameters like policy inception date. An interactive User Interface (UI) will facilitate users in validating and suggesting possible changes to the rules. Users can even contribute their own business rules in natural language, and Gen AI models will seamlessly convert these instructions into executable code in Spark, Python or Structured Query Language (SQL).

1McKinsey & Company
Elevated Data Observability

A pivotal facet of DataOps, data observability enables automated monitoring, data lineage, root cause analysis and data health insights to proactively detect, address, rectify and prevent data anomalies. With Gen AI, data observability will experience a paradigm shift, particularly in automated monitoring, providing critical insights into the health of data and enabling users to identify and resolve issues rapidly and effectively. The data observability workflows will become less complex and more intelligent, ensuring comprehensive monitoring and troubleshooting.

Figure 1: Improving Data Quality with Generative AI

Other instances of Gen AI in data engineering include:

1. **Self-serve Business Intelligence (BI)**
   - Business users can pose scenario-based complex questions to the BI engine in their native language, bypassing the intricacies of SQL / Multi-dimensional Expressions (MDX). This capability yields valuable insights crucial for pivotal business decisions.
   - The effect will be a significant augmentation of the data democratization process within the organization, fostering a reliance on self-serve BI for informed decision-making.
   - As we know, the data democratization process substantially improves the availability and consumption of data by businesses.

2. **Streamlined Data Acquisition and Transformation**
   - With Gen AI models adept at creating Python and SQL scripts, the newer tools and technologies will have advanced transformation defined in their frameworks.
   - Developers can thus employ drag-and-drop transformation, select properties / parameters and use them for specific data operations, devoid of the arduous task of writing complex custom code, and debugging and optimizing it.
   - This shift will substantially reduce the need for custom coding, allowing developers to focus more on solving complex requirements.

3. **Elevated Data Observability**
   - A pivotal facet of DataOps, data observability enables automated monitoring, data lineage, root cause analysis and data health insights to proactively detect, address, rectify and prevent data anomalies. With Gen AI, data observability will experience a paradigm shift, particularly in automated monitoring, providing critical insights into the health of data and enabling users to identify and resolve issues rapidly and effectively. The data observability workflows will become less complex and more intelligent, ensuring comprehensive monitoring and troubleshooting.
To generate more accurate data quality rules, we require data encompassing all possible scenarios. This will allow Gen AI models to simulate and create an exhaustive range of rules to overcome data anomalies and outliers. Synthetic data provides a solution to this quandary. Artificially generated through algorithms, synthetic data replaces real-world events and serves as a substitute for operational data sets, mainly used for validating mathematical models and training AI models. Synthetic data produced by Gen AI models ensures balanced and diverse data with underlying patterns and relationships between data sets, resulting in significantly improved model performance. This facilitates the identification of anomalies such as duplicates, standardization issues, missing values and missing relationships that might otherwise impair data quality.
AI has made remarkable strides in Master Data Management (MDM), enabling features like master data discovery, domain identification, lineage mapping, product classification / categorization, standardization, match / merge and graph-based cross-domain relationships. For example, in the retail sector, product taxonomy has emerged as a pivotal tool, facilitating the logical organization of products through hierarchical structures. The result? Enhanced navigation, improved searchability and a seamless user experience – all of which significantly impact sales. Gen AI algorithms are poised to revolutionize product taxonomy by offering automated solutions for hierarchical structuring, providing invaluable assistance to businesses.

Looking ahead, Gen AI will further optimize these features, bringing increased flexibility. A notable enhancement will be in the identification, consolidation and creation of golden / hybrid records for low-scoring data entries. Gen AI will streamline this process by automatically identifying duplicate master records, clustering them into groups and recommending consolidation methods to create a hybrid golden record that aggregates information from all relevant records. This derived record will satisfy the uniqueness, completeness and other parameters necessary to create a golden record.

Data governance is another area deeply influenced by AI. AI-enabled platforms have advanced active metadata management, cataloging, data asset discovery and monitoring, automated lineage tracking, Role-based Access Controls (RBAC) and regulatory compliance. Let’s explore some areas where Gen AI impacts data governance most profoundly and will see new technological advancements:

1. **AI-enabled ingestion framework** tracks source changes and notifies stewards of any alterations in source systems, automatically adjusting pipelines to lessen the steward burden in pipeline management.
2. **Active catalog** monitors all assets for integrity and quality, alerting stewards of discrepancies.
3. **AI aids in classifying, documenting context, tagging and certifying data assets**, facilitating seamless asset discovery for finding, exploring and consuming data.
4. **Data asset lineage** is traced from source to consumption.
5. **AI is utilized to classify data and tag sensitive information**, with lineage tracking ensuring the traceability of this sensitive data.
Embracing the Future

As data and analytics become increasingly democratized, the potential for Gen AI-driven opportunities in data engineering is limitless. Gen AI will facilitate significant changes in many of these features in the coming days, optimizing existing capabilities and introducing greater flexibility and automation. The future will see new scenarios emerge, causing disruptions in traditional services as new and established players align their offerings based on Gen AI. In this imminent landscape, data experts will find invaluable support in managing their day-to-day ad-hoc tasks, thanks to the adoption of Gen AI-enabled intelligent technologies. With these cutting-edge tools, data experts will be empowered to construct, test, maintain and optimize data services like never before.

To learn how WNS is helping global enterprises harness the power of Gen AI to drive data-led growth, talk to our experts.
About WNS

WNS (Holdings) Limited (NYSE: WNS) is a leading Business Process Management (BPM) company. WNS combines deep industry knowledge with technology, analytics, and process expertise to co-create innovative, digitally led transformational solutions with over 400 clients across various industries. WNS delivers an entire spectrum of BPM solutions including industry-specific offerings, customer experience services, finance and accounting, human resources, procurement, and research and analytics to re-imagine the digital future of businesses. As of June 30, 2023, WNS had 59,871 professionals across 66 delivery centers worldwide including facilities in Canada, China, Costa Rica, India, Malaysia, the Philippines, Poland, Romania, South Africa, Sri Lanka, Turkey, the United Kingdom, and the United States.

About WNS Triage

WNS Triage powers business growth and innovation for 200+ global companies with Artificial Intelligence (AI), Analytics, Data and Research. Driven by a specialized team of over 6000 analysts, data scientists and domain experts, WNS Triage helps translate data into actionable insights for impactful decision-making. Built on the pillars of consulting (Triange Consult), future-ready platforms (Triange NxT), and domain and technology (Triange CoE), WNS Triage seamlessly blends strategy, industry-specific nuances, AI and Machine Learning (ML) operations, and intelligent cloud platforms.

Driving a futuristic edge are WNS Triage’s modular cloud-based platforms and solutions leveraging advanced AI and ML to provide end-to-end integration and processing of data to actionable insights. WNS Triage leverages the combined strength of WNS’ domain expertise, co-creation labs, strategic partnerships and outcome-based engagement models.

To know more, write to us at marketing@wns.com or visit us at www.wns.com

Copyright © 2023 WNS (Holdings) Ltd. All rights reserved.