Deloitte.
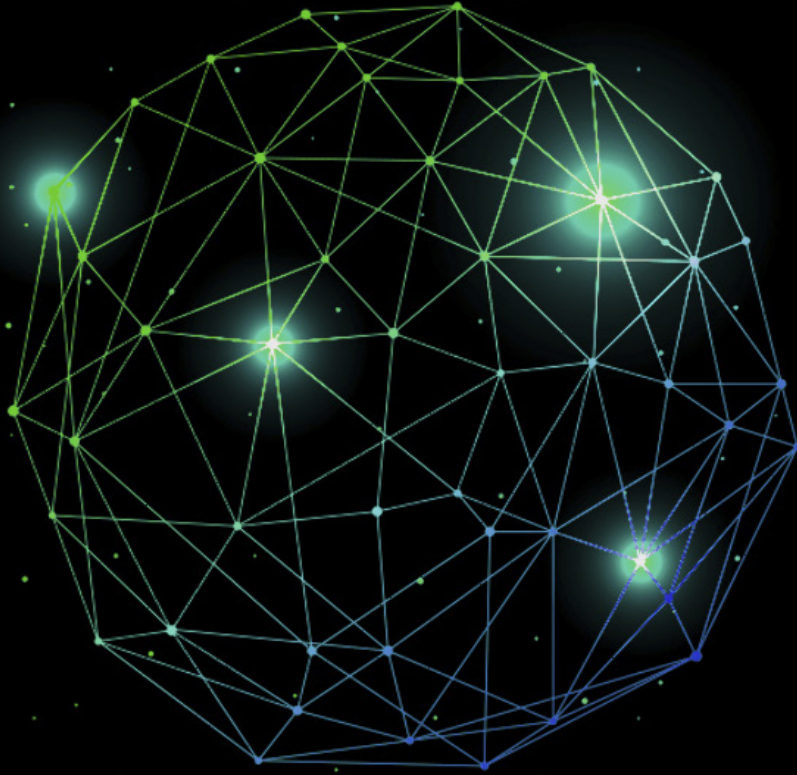
Elevating data standards
in procurement

Unlocking the potential of Generative AI—a thought exercise for use cases

Modern procurement technologies are constantly enhancing their capabilities by leveraging advancements in AI. From there, applications of Generative AI continue to be the discussion topic for procurement leaders—and there are many high-impact use cases to pilot next-gen abilities in supply chain.[1]

However, according to findings from our 2024 Deloitte global survey of 100 chief procurement officers: "While 92% of CPOs are beginning to envision the possibilities of this new technology and plan to invest in it, only 37% were piloting or deploying Generative AI in procurement at the time of the survey. CPOs are still assessing the risk-reward tradeoffs as they begin deploying Generative AI."[2]

One of the biggest obstacles to success, and a major internal barrier to AI in procurement, noted by the surveyed leaders was data quality.

Effective sourcing strategies, supplier relationship management, contract compliance and risk assessment are just some examples of functions that are heavily dependent on accurate and consistent data, particularly with AI-based predictive modeling and scenario analysis.

AI-based intelligent analytics can provide meaningful inputs for key decisions only if the underlying data is accurate and comprehensive. Poor-quality data will likely lead to flawed recommendations and suboptimal decision-making in procurement processes from even the most advanced AI-based capabilities. For example, AI-based scenario modeling to identify risk in supply continuity depends on accurate supplier data, inventory and categorized spend data.

Data quality is not a new topic; but *now* is the time to build its foundation to successfully prepare for the advancements of Generative AI to help enhance end-to-end procurement operations.

Start with effective content management

Procurement operates on a variety of data and generates vast quantities of it—master data such as catalog items, units of measure, supplier information and transactional data such as contracts, purchase orders and invoices.

It is a persistent challenge to maintain quality, which can lead to missing, duplicate or inconsistent data attributes. These data-quality issues are typically caused by limited data governance and/or reliance on manual inputs or disparate systems. And the impacts are substantial, from inaccurate reporting to ineffective decision-making.

Effective content management: get started with data today

Success with AI requires effective data strategy

1 Suppliers
Enable more suppliers to allow users to shop for the best price. Enable more suppliers with catalogs to enhance user experience and expand usability of the platform.

2 Catalogs
Load clean content for catalogs (hosted or punchout) to drive user adoption with an improved guided buying experience, e.g., clear item naming convention, labels and item images.

3 Contracts
Broaden spend managed under contracts to improve spend visibility, drive compliance, and improve downstream invoice efficiency.

4 Master data
Establish a sustainable master data strategy to support upfront cleanup and ongoing data maintenance. Prioritize fixing gaps in master data that support process execution, user experience and reporting.

5 Reporting
Focus on larger business needs for reporting and analytics, and understand the necessary data to derive actionable insights (e.g., spend taxonomy that allows you to analyze cost across categories).

Invest in data quality

This is one area where Generative AI can be transformational, as deep-learning models such as large language models (LLMs) hold the potential to significantly enhance data quality. In this article we outline about some of the high-impact data enhancements for procurement that can benefit from Generative AI. But first, let's look at the key aspects of enhancing the data quality.

Data normalization: Generative AI can automate the data cleansing process by identifying and correcting errors in datasets. LLMs can be trained on procurement-specific terminology and industry data to recognize and standardize inconsistent entries, and to identify and remove duplicative entries even when exact matches are not present.

Data imputation: Generative AI can address the problem of missing data and generate plausible substitutes for missing values based on existing patterns. Models like generative adversarial networks (GANs) and transformer-based architectures can be trained on existing data to learn complex patterns and correlations.

Data augmentation: Generative AI with models like variational autoencoders (VAEs) can augment procurement data by creating synthetic datasets to enrich data. This is also particularly useful for scenario analysis—for example, simulating the impact of a sudden supplier outage on costs and timelines, which allow an organization to develop contingency plans.[3]

Common master data challenges in procurement

Category taxonomy
- Category taxonomy not used consistently across systems
- Governance process for changes to taxonomy not defined
- Legacy transactional data tied to existing commodity codes

Conversions
- Lack of strategy on whether legacy transactional data should be converted, and a typical recommendation to only convert open POs during cutover

Spend data
- Initial plan for determining approach to historical data not defined (i.e., cleanse normalize and load, or build as you go)

Supplier
- Duplicate supplier data
- Outdated contact info
- Missing supplier data: email, relationship owner, contracts (for catalog suppliers)

User data
- Incomplete or inconsistent human resources (HR) data across systems
- "Levels" not part of HR record or inconsistent by system, business unit or geography

Accounting data
- General ledger (GL) and commodity table mapping is missing or not accurate
- Old GL accounts still active, requiring cleanup

Products / Materials
- Pricing information missing for items
- Item-level descriptions are not consistent, hampering effective catalog search
- Duplicate items

Units of measure (UOM)
- UOM use inconsistent across systems
- Level of effort to normalize UOM, if required, is underestimated
- UOM is not mapped to United Nations Standard Products and Services code, leading to errors

HOW TO ENABLE SUCCESS
- To begin: Assess quality of master data
- Prioritize master data gaps that could negatively influence functionality and user experience, e.g., suppliers, catalog items
- Use leading practices to fix highest-value areas, e.g., data affecting spend visibility and reporting, catalogs, access to suppliers
- Develop a realistic plan to clean and enrich data; include the team that will work on it and a governance plan for ongoing master data management

Let's look at some of the key use cases for enhancing procurement data quality through AI-powered data normalization, imputation and augmentation. These specific scenarios are a thought experiment in how Generative AI abilities can be harnessed to further enhance procurement data.

Supplier and item master data enhancement: Supplier data is often fragmented and inconsistent, especially when there are multiple systems in use across different entities or regions. Generative AI can consolidate and enrich supplier profiles by synthesizing supplier data from multiple sources including external databases. For instance, there are straightforward corrections where generative models can detect inconsistent supplier names and automatically standardize them by leveraging external data bases; and more complex corrections where parent-child entities can be identified and linked. Accurate and comprehensive supplier profiles can help make risk assessment models more beneficial, even for companies leveraging the latest solutions for supplier risk assessment.

Another high-impact data element is catalog items. While many categories can leverage punchouts where suppliers maintain this data, services catalogs or internally hosted items are still relevant for many purchasing scenarios. The accuracy and consistency of this data is crucial for effective ordering and downstream processes. If catalogs are missing prices or units of measure, a generative model can predict the likely values based on historical purchase data, supplier trends and market conditions.

Intelligent categorization and classification: Spend data classification has been a pain point for many large organizations. Accurate classification of procurement data is crucial for effective spend analysis and strategic category management. Generative AI models utilizing natural language processing (NLP) can classify large subsets of master data such as suppliers or item / service descriptions against appropriate spend categories. This classification is equally if not more relevant to transactional data, such as contracts or purchase orders, to make meaningful decisions based on spending patterns.

A more fundamental issue for many organizations remains "What is the right category structure for my industry?" Today's AI models can also more intelligently analyze sourcing patterns and spend data to recommend an effective category structure that is important for procurement strategies. Generative AI, especially, holds the potential to produce higher quality outputs compared to machine learning-based models leveraged in the past.

Challenges and considerations: While the applications of Generative AI in enhancing data quality are promising, there are challenges to consider related to data privacy and large training data requirements, as well as interactions with legacy systems. Robust data governance policies and investments in robust architectures will be required to leverage these models and continue to enhance enterprise digital procurement capabilities.

▶

Moving forward

Now is the time to invest in the foundation to be ready to leverage the Generative AI revolution. Procurement leaders can approach data quality through sophisticated AI tools for data cleansing, imputation, classification and augmentation. By addressing common data-quality issues and providing more accurate, comprehensive insights, Generative AI can empower procurement teams to make better decisions, reduce risks and drive efficiency. As adoption of these technologies grows, we can expect a new era of data-driven procurement that delivers strategic value across the supply chain. Begin your journey by evaluating your current data issues and establishing robust data governance models–assess solutions in the market leveraging AI powered data management.

Authors

Janet Lu
Consulting Managing Director
Deloitte Consulting, LLP
jalu@deloitte.com

Megha Chaudhary
DC Senior Manager
Deloitte Consulting, LLP
mechaudhary@deloitte.com

Endnotes

1. Ryan Flynn, Mike Deng, Vinay Rajani and Ayushman Kaul, "2024 Global CPO GenAI survey shows how CPOs are approaching AI in procurement and sourcing," Deloitte, August 21, 2024, https://www2.deloitte.com/us/en/blog/business-operations-room-blog/2024/generative-ai-in-procurement.html, accessed December 3, 2024.
2. Flynn et al, "2024 Global CPO GenAI survey."
3. Spencer Young and Phill Domschke, "Start by using AI to clean up your data," Deloitte, June 11, 2024, https://www2.deloitte.com/us/en/blog/business-operations-room-blog/2024/ai-in-data-analysis.html, accessed December 3, 2024.

Deloitte.