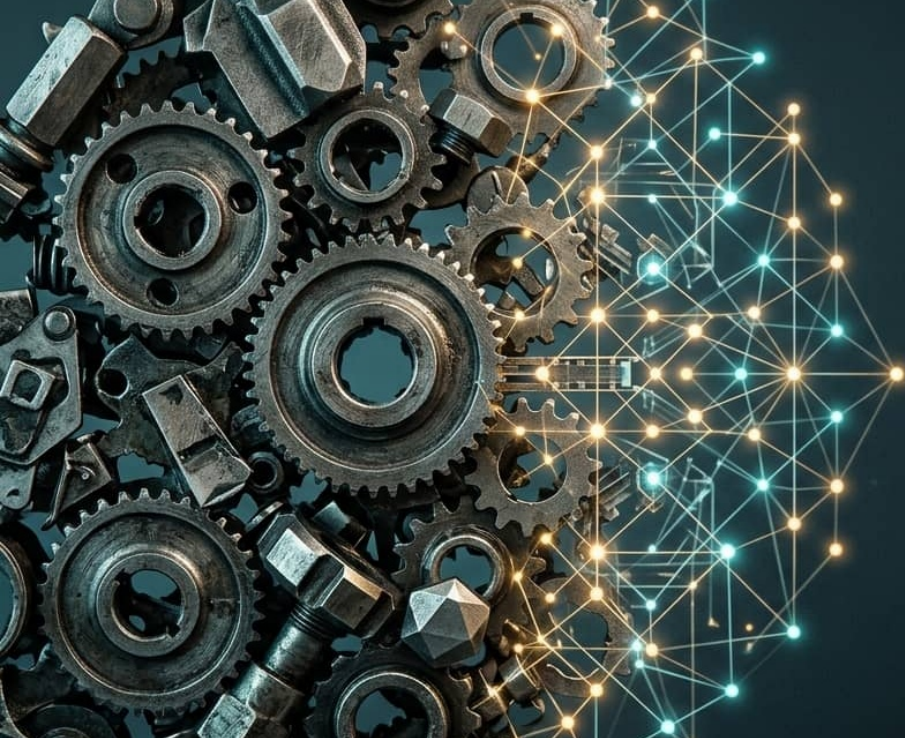




Mapping multilingual spend data with AI techniques

Mapping multilingual spend data requires moving beyond literal machine translation to establishing unified semantic data structures across enterprise systems. By combining deep learning with domain-specific knowledge engineering, global organizations can translate, categorize, and normalize material master data in real time. What the market still calls "mapping" is a static, manual, point-to-point process. Creactives replaces this legacy approach with an Enterprise Data Space (EDS): a network of individual semantic digital twins (for materials, services, suppliers, and assets) connected to one another, which together normalize and harmonize data automatically. This semantic foundation enables autonomous AI agents to analyze spend visibility, consolidate supplier contracts, and enforce purchasing compliance across fragmented enterprise resource planning (ERP) systems, regardless of the original language used at the plant level.

When procurement teams operate across multiple continents, disparate terminology often masks massive savings opportunities. A factory in Germany might purchase a replacement bearing under the description "Kugellager 6204 2RS", while a facility in Mexico enters "Rodamiento rígido de bolas

6204", and a site in the United States logs "Ball bearing 6204-2RS". Without an intelligent system connecting these records, your ERP views them as three entirely distinct items. As a result, you lose regional purchasing leverage, accumulate duplicate inventory, and miss compliance leaks.

Why literal translation fails for global procurement records

Standard translation engines fail when applied to procurement data because industrial descriptions rely heavily on shorthand, regional abbreviations, manufacturer part numbers, and unstructured technical parameters. General-purpose large language models or consumer translation APIs translate words rather than technical intent. If an algorithm translates an abbreviated German description like "Flansch m. Nut" literally, it might miss the exact engineering standard, preventing the item from matching an existing category code in your central taxonomy.

Procurement data is also notoriously inconsistent. Over decades of local operations, plant engineers and purchasing clerks develop idiosyncratic logging habits. You will find entries missing manufacturer names,

descriptions where dimensions are buried inside product titles, and vendor codes duplicated under slightly different regional entities.

Machine translation cannot solve this problem on its own because translation is not classification. While legacy approaches attempt point-to-point spend data mapping, resolving multilingual procurement requires you to extract technical attributes, identify product categories, and reconcile naming conventions simultaneously. To achieve accurate procurement spend visibility, you need an architecture that understands industrial context at the entity level.

How AI replaces manual spend data mapping

Replacing manual mapping with AI requires a hybrid approach that combines deep learning, natural language processing, and symbolic artificial intelligence.

Deep learning models excel at handling language variation, interpreting messy text, and inferring context from incomplete records. However, deep learning models can generate hallucinations or lack deterministic precision when handling strict part specifications. Knowledge engineering, on the other hand, provides structured ontologies, deterministic rules, and hierarchical taxonomies (such as UNSPSC, eCI@ss, or custom corporate classification trees).

A digital twin is a virtual, mathematically precise model of one specific class of enterprise object: a material, a supplier, an asset, or a service. Creatives builds one for each domain, then connects them semantically to one another. That connected sum is the Enterprise Data Space: a single space where every material, supplier, and asset is uniquely identified and interconnected.

Specialized platforms like Creatives use proprietary AI to construct these individual digital twins, normalizing disparate text across dozens of languages. Instead of translating descriptions into English and then attempting to run category rules, the platform connects the raw multilingual text directly to universal semantic concepts in the knowledge graph. The system recognizes that whether an item is

logged in German, Japanese, Portuguese, or Italian, it refers to the exact same physical component with identical operating characteristics.

By establishing this connected layer, you transform passive database entries into dynamic assets. Organizations using AI for procurement spend visibility and analytics leverage this semantic layer to eliminate data blind spots, consolidate tail spend, and monitor global procurement trends without relying on manual data-cleaning projects.

Unified data space vs traditional data lakes for procurement

For years, enterprise IT departments attempted to resolve spend fragmentation by funneling all local ERP extracts into a central data lake. While data lakes are effective for storing vast quantities of raw information, they consistently fall short for procurement spend analytics.

Data lakes rely on a "schema-on-read" philosophy. They ingest raw, messy tables from SAP, Oracle, and legacy local systems without harmonizing the data during ingestion. When business analysts or AI agents attempt to query the data lake, they must reconstruct context on the fly. Because different plants use different part numbers, units of measure, and local descriptions, queries across the data lake return fractured, incomplete results.

An Enterprise Data Space (EDS), historically referred to as a unified data space, establishes an interconnected semantic layer on top of your transactional systems. Rather than treating records as isolated database rows, the EDS links domain-specific digital twins into a shared knowledge graph. This foundation allows teams to stop querying records and start navigating relationships. The EDS leaves the underlying operational records intact while creating a shared, contextualized intelligence fabric. Table 1 sets the two architectures side by side.

When you build an Enterprise Data Space, you do not need to replace every legacy ERP across your business units immediately. Instead, the semantic architecture federates information from multiple environments, translating and structuring material records so

that procurement teams view spend through a single pane of glass.

How autonomous AI agents drive procurement spend visibility

Once a unified data structure is active, autonomous AI agents can perform cross-contextual work that human category managers cannot execute at scale. However, the connection between data foundation quality and agentic capability is direct: without a unified semantic layer, even the most advanced AI agents cannot accurately interpret multilingual data, analyze spend, or enforce compliance. Grounded in an Enterprise Data Space, these vertical procurement agents analyze transactions across subsidiaries, identifying anomalies and recommending commercial actions in real time.

Autonomous agents deliver tangible value in three primary operational areas:

1. Cross-plant duplicate identification

Without a semantic layer, duplicate materials hide under local part numbers across global plants. An AI agent operating across connected digital twins identifies identical replacement parts across facilities, even when descriptions use different languages and distinct catalog codes.

By knowing how to find duplicate materials with AI in global inventory, organizations can avoid redundant safety stock purchases, reallocate dormant spares across regions, and cut warehouse holding costs.

2. Tail spend rationalization

Uncategorized tail spend accounts for significant capital leakage in global enterprises. AI agents parse unstructured free-text purchase orders, categorize them into specific sub-tiers in your taxonomy, and identify unmanaged spend patterns. If three separate international facilities are buying the same cleaning solvents from regional distributors at divergent rates, the agent flags the discrepancy and prompts category leaders to negotiate a consolidated contract.

3. Real-time master data governance

Rather than correcting material master errors months after purchase orders are issued, autonomous agents intervene at the requisition stage. As an engineer begins typing an item request in their native language, the agent interprets the technical attributes, matches them against the global master catalog, and suggests an existing catalog item. This prevents the creation of duplicate master records before they enter the ERP.

How to compare spend analytics vendors for multi-language data

Evaluating spend analytics software requires looking beyond attractive visualization dashboards. The true test of a procurement platform lies in how it handles dirty, multi-language data at the data-foundation level.

When comparing vendors, assess how their solutions perform across these critical dimensions:

Natural language architecture: translation vs semantic ontologies

Many conventional spend analytics tools rely on third-party translation APIs layered over simple keyword-matching scripts. These solutions struggle with complex industrial terminology, metric-to-imperial conversions, and regional slang.

Look for tools that offer semantic interconnection for procurement, utilizing built-in ontologies developed specifically for materials management, maintenance, repair, and operations (MRO), and complex direct spend. The system should understand that two completely different string descriptions represent the same physical entity without needing a human to manually write matching rules.

Continuous master data governance vs periodic data cleansing

Traditional consulting vendors offer episodic "data-cleansing projects" that clean your spend records once a year. Within three months of project completion, new dirty records accumulate, and the data degrades again.

Leading platforms provide continuous governance. They clean historical records to establish a baseline, then operate inside your live transactional systems to normalize and classify new requisitions as they happen.

Speed of integration during mergers and acquisitions

When acquiring new businesses, integrating their legacy ERP data into your central reporting structure can take years. Platforms powered by specialized AI can harmonize disparate master data up to twelve times faster during acquisitions, rapidly identifying overlapping suppliers, duplicate spare parts, and immediate volume discount opportunities.

Preparing master data for global ERP consolidation

Clean, multilingual master data is not just an operational convenience; it is a prerequisite for enterprise digital transformations, including migrations to SAP S/4HANA. Migrating unstandardized, multilingual records into a new ERP system replicates legacy inefficiencies on expensive modern infrastructure.

By replacing manual spend data mapping with AI-driven semantic harmonization before and during migration, you ensure that:

- Every material record contains structured technical attributes (voltage, thread size, material composition) rather than ambiguous free text.
- Parent-child supplier relationships are resolved globally, exposing vendor fragmentation across regional operating units.
- Material records are linked to existing equipment asset hierarchies, clarifying which parts support critical plant operations.

Establishing an Enterprise Data Space eliminates the need for manual, point-to-point mapping and bridges the gap between local plant operations and corporate procurement strategy. When your procurement AI agents and human analysts share a single, multilingual understanding of every dollar spent, your organization unlocks true spend visibility, reduced inventory risk, and genuine global purchasing power.

Table 1: Traditional data lake vs Enterprise Data Space

Feature	Traditional data lake	Enterprise Data Space (EDS)
Data ingestion	Dumps raw, multilingual data without normalization	Ingests and harmonizes data within a common semantic ontology
Multilingual handling	Requires external batch translation before query	Real-time multilingual normalization at ingestion
Contextual relationships	Stores tables in isolation; relationships must be manually joined	Maintains a knowledge graph linking materials, suppliers, and assets
Data maintenance	Periodic, high-cost manual data-cleansing cycles	Continuous, autonomous governance and duplicate detection
AI agent readiness	Low; agents stumble over dirty descriptions and naming mismatches	High; autonomous agents navigate standardized entities directly

Creactives builds a trusted Data Foundation that creates a Digital Twin of your Master and Spend Data. Our AI-powered solutions eliminate errors, optimize legacy systems, and unlock a wide range of use cases across Procurement & Supply Chain.



SPEND ANALYTICS

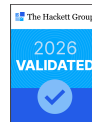


**MATERIALS & SERVICES
MASTER DATA GOVERNANCE**



**PROCUREMENT GUIDANCE
AUTOMATION**

CREACTIVES: SOLUTIONS SELECTED OVER YEARS BY OPINION LEADERS



"Creactives is a best-of-breed tool that offers standard spend analytics features, such as data cleansing and analysis. In addition, it specializes in materials MDM and is built for automated parts and material classification."

Abigail Ommen, Research Analyst & Analyst Production Manager