Enterprise spend analytics for fragmented data: a guide

Effective **enterprise spend analytics for fragmented data** requires a transition from manual, rule-based categorization to AI-driven semantic harmonization. By leveraging deep learning, global organizations can unify disparate material and supplier records across multiple languages to identify hidden stock and eliminate procurement duplicates. This approach provides visibility that legacy master data management (MDM) tools cannot achieve, creating a resilient data foundation for strategic sourcing.

The logical bridge: why general master data quality dictates spend visibility

You cannot aggregate and analyze spend across a global enterprise if the underlying material records are semantically fragmented. Spend analytics is only as good as the semantic layer beneath it; therefore, the quality of **general master data**, **encompassing both materials and suppliers**, is the absolute precondition for meaningful insights. When records are inconsistent, the result is "spend blindness," a state in which procurement leadership cannot see total requirement volumes because the underlying data lacks a common language.

Why fragmented data hinders global supply chain visibility

For most multinational corporations, the primary obstacle to supply chain visibility is a lack of semantic connection between systems. Procurement information is often siloed, with materials and assets described in unstructured free text. When data is fragmented, a supply chain director cannot easily see that a **"Coussinet à billes SKF 6204"** in a French plant is the same item as a **"skf brg 6204"** in a USA facility.

This fragmentation leads to "hidden stock"—inventory that exists within the organization but remains invisible to procurement teams due to inconsistent naming conventions. Without a unified view, companies overorder materials they already have, miss volume discounts, and incur unnecessary inventory costs. Achieving true visibility requires **enterprise spend analytics for fragmented data** that understands the technical intent of every record beyond surface-level text.

Legacy MDM vs. deep learning: identifying the visibility gap

Traditional master data management (MDM) tools, such as SAP MDG, were designed to govern structured fields like vendor names and tax codes.

These domains have limited, well-defined values managed via manual workflows. However, the material master domain is fundamentally different due to its technical complexity and massive scale.

The limitation of rule-based material master data governance
Material master data is unstructured and often encompasses millions of records. Legacy tools rely on human-defined rules and "fuzzy logic" to match records. For an industrial enterprise, manually configuring these rules is a multi-year effort that requires extensive maintenance because data changes necessitate new validations that pile up on the older ones.

Standard **SAP MDG material master** implementations provide a robust governance framework but lack the inherent intelligence to understand material substance. They can enforce field completion, but they cannot determine if the content is a duplicate of an existing item described differently.

How deep learning enables semantic duplicate detection
Deep learning shifts the focus from word matching to meaning matching. Specialized AI platforms, such as those developed by Creactives, use knowledge engineering to analyze technical attributes in a description.

This **semantic duplicate detection** allows the system to recognize that "Ruptor 400V Siemens 5SY4310/7" and "Circuit breaker 5SY4310-7" are identical, even if the phrasing differs. By identifying these matches, organizations can reduce their material master records by 20% or more, directly lowering inventory value.

Harmonizing multilingual procurement data for strategic sourcing
Global procurement teams face complexity originating at the design phase within PLM (product lifecycle management) and BOM management systems. Material descriptions are typically written by local engineers in native tongues and technical shorthand, making multilingual procurement data harmonization a core requirement for resilient operations.

Bridging the language gap with native AI
Most spend analytics vendors attempt to solve language barriers through machine translation, which often fails with technical jargon. A more effective approach is using AI trained natively on industrial taxonomies across 25+ languages.

Creactives' technology does not just translate; it categorizes and extracts attributes directly from the source language. This creates a digital twin of the material, ensuring that technical specifications are standardized, regardless of whether the input was in Portuguese, Chinese, or English. This level of harmonization allows CPOs to aggregate spend across dozens of countries to negotiate better global contracts.

Achieving a competitive edge with a unified data space (UDS)
Modern **material master data governance** aims to create a digital twin for every procurement object. This semantic representation forms a unified data space (UDS), serving as the foundation for spend analytics and AI agents. When data is semantically connected within a unified data space:

- **Strategic sourcing is automated**: AI agents instantly group fragmented spend into the correct categories.
- **Inventory is optimized**: Redundant stock is identified across different regions through a single source of truth.
- **Compliance is enforced**: Maverick spend is reduced because users can easily find existing items.

Unlike legacy bulk-cleansing tools that require constant manual stewardship, deep learning models evolve through user interactions, ensuring data quality improves without periodic manual cleaning projects.

Implementing real-time governance: the smart creation approach
Focusing solely on cleaning historical data without protecting the future is a common error. Real-time governance, or "smart creation," integrates AI directly into the material request process to prevent poor-quality information from entering the ERP.

As a user describes a new part, the AI:

1. Searches the global database for existing duplicates across all languages.
2. Performs **parallel categorization**, suggesting the correct classes across multiple taxonomies simultaneously—including custom proprietary structures and international standards like UNSPSC or eCl@ss—without manual mapping.
3. Automatically normalizes the description and extracts technical attributes.

This front-end prevention can reduce new material creation by up to 35%, significantly slowing the growth of the material master.

Case study: driving ROI through material standardization

A global energy and utilities multinational operating in 23 countries and holding 2 million material records faced a massive harmonization challenge. With eight different ERP systems and eight languages, they lacked a consistent taxonomy. By deploying Creactives as a material domain specialist alongside their existing SAP MDG, the company achieved:

- **95% data accuracy** in historical spend.
- **$30 million+** in identified duplicate stock value.
- **20% reduction** in total material master records.

This "proven co-existence" model allows **SAP** MDG to own the governance workflow while Creactives provides the semantic intelligence the legacy system lacks.

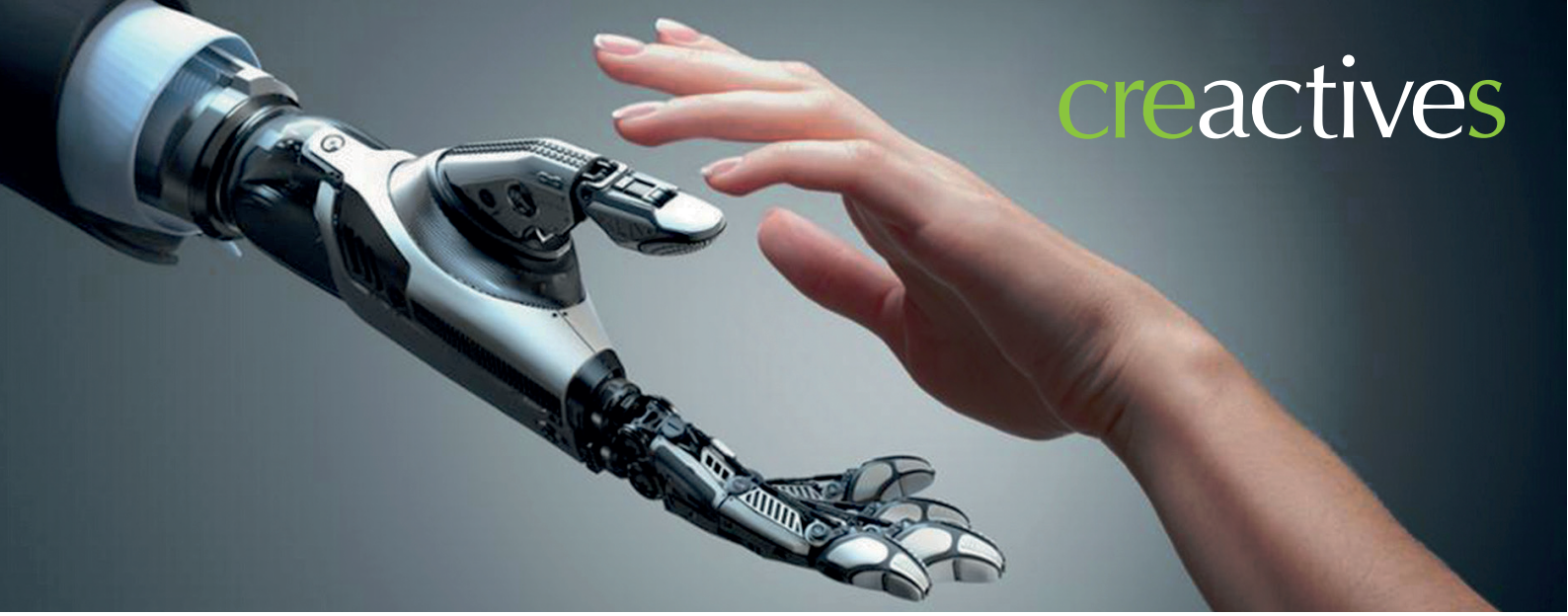
Choosing the right platform for your enterprise data foundation

Supply chain directors should prioritize tools that offer out-of-the-box industrial intelligence rather than just a software framework.

For organizations migrating to **SAP** S/4HANA, the strategy should be to "clean before you migrate." Deploying a deep learning solution on legacy systems ensures you deliver a deduplicated, enriched material master to the new environment.

To learn more about how to unify your fragmented procurement data and gain a competitive edge through semantic AI, explore **Creactives' solutions for master data quality.**

Feature	Legacy MDM / ERP tools	Creactives Technology
Taxonomy	Manual configuration and mapping required	**Parallel categorization** across custom and standard (UNSPSC, eCl@ss) taxonomies
Language	Primarily language-neutral (IDs/codes)	Native support for 25+ languages
Duplicate detection	Lexical/fuzzy (word-based)	Semantic (meaning-based)
Governance	Workflow-centric	Intelligence-centric (smart creation)
Implementation	Months/years of rule-mining	Business value from day one



creactives

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

Creactives Group SPA
Via Fermi, 4 – 37135 Verona – Italy

creactives

creactives.com
linkedin.com/company/creactives-spa