



Guide to identifying multilingual duplicate spare parts

Successfully identifying multilingual duplicate spare parts across distributed facilities requires decoupling engineering specifications from local text descriptions. By extracting physical attributes, standardizing units of measure, and mapping records to a centralized semantic ontology, industrial enterprises can uncover identical MRO items masked by localized jargon. This systematic approach eliminates redundant safety stock, slashes working capital, and prevents unplanned downtime across European and Latin American manufacturing operations.

Multinational manufacturers running plants across Germany, France, Spain, Brazil, and Mexico face a recurring operational challenge. While maintenance engineers at different sites maintain identical turbines, compressors, and packaging lines, their enterprise resource planning (ERP) systems treat identical spare parts as completely different inventory items. When maintenance teams do not know that a replacement motor sits idle in a regional warehouse three hours away, they place costly emergency purchase orders with external suppliers, driving up MRO inventory duplicates and tied-up capital.

Why identical critical spare parts become invisible across international operations

The invisibility of identical components stems from fragmented systems, localized naming habits, and language barriers. Industrial networks typically expand through international acquisitions, leaving enterprises with separate ERP instances such as SAP, Oracle, and local legacy systems.

Plants in European countries like Germany, Italy, and France catalog items in their native languages, using local engineering shorthand. Meanwhile, facilities in Mexico, Colombia, and Brazil record the same components in Spanish and Portuguese. Even within the same language, regional terminology diverges: a Spanish plant might log an item as an *empuñadura*, while a Mexican facility logs it as an *agarre*, and an Argentine site labels it a *manija*.

Local maintenance staff also introduce idiosyncratic abbreviations. One technician writes *mot trif 400v 5kw*, another logs *Mtr 3F 5.0kW 400 Volt*, and a European counterpart logs *Drehstrommotor 5 kW 400V*. To traditional ERP search engines, these three descriptions share zero lexical overlap.

Because standard ERP systems search via exact text strings, they cannot recognize that these entries describe the exact same physical component. As a result, critical spare parts are purchased repeatedly across business units, inflating storage costs while critical machines remain at risk of extended outages.

The failure of literal translation and keyword matching

When procurement leaders attempt to resolve this visibility gap, their first instinct is often to deploy literal machine translation or rule-based string matching. These approaches fail in industrial contexts because they treat MRO catalog entries like regular conversational text.

Generic machine translation tools do not understand industrial context, technical abbreviations, or noun-adjective positioning across Romance and Germanic languages. For example:

- A German catalog might list a *Rillenkugellager 6205-2RSH SKF*.
- A French catalog lists *Roulement rigide à billes SKF 6205 2RSH*.
- A Spanish catalog records *Rodamiento rígido de bolas 6205.2RSH marca SKF*.
- A Brazilian Portuguese catalog inputs *Rolamento rígido de esferas 6205-2RSH – SKF*.

While a skilled reliability engineer recognizes these as the exact same deep-groove ball bearing, automated translation engines often distort technical terms. Translating an Italian entry like *valvola a farfalla* into Spanish might yield a literal reference to an insect rather than a butterfly valve, while German compound nouns (*Druckbegrenzungsventil*) fragment into nonsensical terms when processed through basic translation APIs.

Furthermore, literal translation does not resolve structural mismatches in data entry. One plant may place the manufacturer part number (MPN) at the beginning of the text field, while another hides it inside an unformatted comment field or omits the manufacturer name entirely. Rule-based tools require clean, predictable inputs, which makes them ineffective against real-world plant

records. Achieving true data coherence requires mapping multilingual spend data with AI techniques, which move beyond literal translation toward deep semantic understanding.

The following four records, drawn from four different plants, all resolve to the same engineering item: a socket head cap screw, M8 thread, 40 mm length, grade 8.8, zinc-plated.

- **Germany (Munich):** *Zylinderschraube DIN 912 M8x40 8.8 verz.* — literally “cylinder screw DIN 912 M8x40 8.8 galvanized”.
- **Spain (Valencia):** *Tornillo cil con hex int M8x40 zinc clase 8.8* — literally “cylindrical screw with internal hex M8x40 zinc class 8.8”.
- **Brazil (Curitiba):** *Parafuso allen cilindrico M8 x 40mm aco 8.8 zinc* — literally “allen cylindrical screw M8 x 40mm steel 8.8 zinc”.
- **Mexico (Monterrey):** *Torn allen C/C M8-1.25 X 40 G8.8 galv* — literally “screw allen C/C M8-1.25 X 40 G8.8 galv”.

Core methodology: semantic multilingual material classification

Resolving duplicates across global plants requires a transition from lexical searching to semantic feature extraction. The objective is to convert unstructured, multi-language text descriptions into language-agnostic, structured engineering records.

Multi-plant local ERP records

DE: Rillenkugellager / ES: Rodamiento / BR: Rolamento



Domain-specific knowledge engineering

- Multilingual industrial terminology parsing
- Unit conversion (imperial / metric normalization)
- Manufacturer and part number (MPN) isolation



Standardized attribute matrix

- Family: bearing > ball bearing > deep groove
- Bore 25 mm | outer diameter 52 mm | width 15 mm
- Seal: rubber contact (2RSH/2RS) | clearance: CN



Unified digital twin

Identical item recognized across all global plants

Decoupling attributes from natural language

Semantic AI systems isolate the functional variables that define an engineering component, regardless of how the surrounding description was phrased. These physical variables include:

- Primary functional category (such as valve, pump, bearing, seal, or relay).
- Dimensional characteristics (such as inner diameter, outer diameter, length, and thread pitch).
- Physical tolerances and material grades (such as AISI 316 stainless steel, nitrile rubber, or bronze).
- Operational ratings (such as operating voltage, pressure capacity, temperature limits, and power output).
- Manufacturer name and manufacturer part number (MPN).

Once these attributes are isolated, the system translates distinct national measurement formats into standardized units. A plant in Mexico logging pipe fittings in inches and a plant in Germany logging metric equivalents are automatically aligned into a single technical representation.

Building active domain knowledge graphs

Rather than relying on generic large language models that generate probabilistic approximations, industrial inventory reconciliation demands deterministic precision. Domain-specific knowledge graphs store relationships between thousands of mechanical, electrical, and pneumatic standards.

When the knowledge graph processes a German record referencing *DIN 912* and a Mexican record referencing *tornillo allen*, it identifies both as ISO 4762 socket head cap screws. It understands that *Viton* and *FKM* designate the same fluoroelastomer material, ensuring that language-specific marketing terms and chemical abbreviations resolve to an identical classification.

Through this methodology, enterprises build an authoritative foundation for finding duplicate materials with AI in global inventory,

cutting inventory holding costs while preserving operational resilience.

Step-by-step field guide to resolving duplicate critical spares

Executing an inventory deduplication initiative across European and Latin American plants requires a phased approach that protects day-to-day operations while progressively harmonizing master data records.

1. **Ingest and centralize multi-ERP catalogs without disruption.** Do not attempt a massive, multi-year ERP migration before establishing inventory visibility. Extract active MRO material masters, purchasing history, and inventory balance records from each European and Latin American plant as they currently exist. Retain local material numbers, plant-specific storage locations, and local language descriptions, and centralize these records in an external semantic data layer where analysis can occur without interfering with ongoing maintenance schedules or warehouse operations.
2. **Normalize and isolate high-value critical components.** Sort inventory items by criticality and annual spend, and focus the initial deduplication sweep on high-consequence machinery components (turbine rotors, multi-stage pump impellers, compressor seals), long lead-time automation assets (programmable logic controllers, variable frequency drives, servomotors), and expensive consumables with high turnover (mechanical seals, specialized bearings, hydraulic valves). By focusing on critical spares first, supply chain leaders achieve immediate return on investment by canceling planned purchase orders for items that already exist in regional or international stock.

3. **Parse attributes through multilingual technical intelligence.** Pass raw text through specialized artificial intelligence trained on international technical taxonomy. The engine translates technical vernacular from German, French, Italian, Spanish, and Portuguese into standardized technical taxonomy; extracts specific numerical values, matching them to their physical dimensions, pressure limits, and electrical loads; and cleanses and validates manufacturer names, standardizing variations like *Rexroth*, *Bosch-Rexroth*, and *Mannesmann Rexroth* into a single canonical entity.
4. **Group records by semantic fingerprint.** With attributes extracted, the system groups items that possess identical engineering parameters into clusters, regardless of their source plant or description text. Each cluster reveals instances where two or more local material codes represent the same physical spare: an emergency pump impeller stored under material code *DE-10492* in Ludwigshafen matches an unassigned spare stored under *MX-88301* in Coatzacoalcas. Maintenance managers can inspect the cluster, verify technical compatibility, and link the records.
5. **Establish cross-plant pooling and rationalized min-max levels.** Once duplicate items are linked, recalculate inventory safety buffers across regional clusters. Instead of three Latin American plants holding two identical backup servomotors each (totaling six redundant units across the region), the organization can establish a shared regional pool of three units, freeing up substantial capital while maintaining uptime guarantees.

Operational differences: European vs. Latin American MRO data

Addressing duplicates across these two continents requires accounting for distinct regulatory, logistical, and data practices that influence catalog structures.

Standard designations and regulatory references

European plants heavily reference international and continental standards

organizations such as DIN (Deutsches Institut für Normung), ISO, and AFNOR. Descriptions often omit descriptive words entirely, relying instead on concise standard strings such as *DIN 7991 M10x30 A2*.

Latin American facilities often blend American standards (ANSI, ASME, ASTM, SAE) with metric and imperial units due to supply chain links with North America. A Mexican facility might purchase an ANSI flange for a pipeline designed in the United States, while a sister plant in Germany runs an identical line built to DIN specifications. The semantic engine must bridge metric and imperial systems to detect functional interchangeability.

Lead time buffers and local purchasing autonomy

Plants in Latin America often face longer import transit times and unpredictable customs clearance processes for components manufactured in Europe or Asia. Consequently, local warehouse managers frequently create localized variants of existing parts codes to facilitate immediate spot-buying from domestic distributors.

These unofficial spot-purchases introduce non-standard supplier part numbers into local ERPs, obscuring the original equipment manufacturer (OEM) details. Reconciling these records requires an engine capable of backtracking distributor part numbers to original engineering blueprints.

Establishing continuous ERP master data harmonization

Deduplicating existing stock delivers immediate working capital relief, but organizations must prevent bad data from re-entering the environment. Without continuous governance, local plants naturally revert to old behaviors, generating new duplicate material master records within months.

Enterprises must embed automated governance at the point of requisition. When an engineer in São Paulo requests a new material code, the software must read the raw, free-text Portuguese description, extract its attributes in real time, and query the entire global database.

If an identical or functionally equivalent component already exists in Monterrey, Barcelona, or Frankfurt, the system flags the match immediately. The engineer can review existing company stock or link their request to the established global material master code instead of creating a duplicate entry.

This continuous data governance approach is critical for organizations preparing for large digital transformations. When moving legacy systems to modern environments like SAP S/4HANA, pre-cleansing multi-language data eliminates the risk of migrating garbage data into the new enterprise core. For organizations dealing with deep visibility gaps across multiple units, deploying enterprise spend analytics for fragmented data provides the clear foundation needed to identify hidden inventory and optimize supply contracts globally.

The strategic value of semantic digital twins in procurement

Solving the multilingual duplicate spare parts problem requires moving beyond flat databases, static spreadsheets, and point-to-point translation scripts. Enterprise leaders achieve sustainable visibility by modeling their entire physical infrastructure through a semantic digital twin.

A semantic digital twin creates a dynamic, machine-legible representation of an organization's industrial reality. It links:

- Materials (physical components, raw materials, consumable goods).
- Equipment and assets (production lines, pumps, gearboxes, operating sites).
- Suppliers and manufacturers (global OEMs, authorized distributors, local vendors).
- Technical attributes (pressure ratings, power limits, dimensions, materials of construction).

Creactives provides enterprise software built specifically around this architecture. Powered by proprietary artificial intelligence that combines deep learning with active knowledge engineering, Creactives builds enriched semantic digital twins of materials, suppliers, and assets. These digital twins connect into

an Enterprise Data Space, unifying multi-ERP procurement data into a single, trusted environment where organizations can navigate relationships across disparate systems instead of querying isolated records.

Operating in real time across any language or alphabet, the platform identifies identical critical spare parts across global plants regardless of terminology, spelling variations, or localized abbreviations. By managing technical attributes automatically and incorporating the practical knowledge of field engineers, Creactives enables global enterprises to eliminate duplicate MRO inventory, accelerate S/4HANA migrations, and secure cross-border supply chain resilience.

Transforming fragmented, multilingual maintenance catalogs into a semantically connected digital ecosystem turns maintenance data from an operational bottleneck into a source of capital efficiency and sustained operational uptime:

<https://www.creactives.com/>

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