



How to clear a five-figure material master backlog in weeks: A step-by-step guide to AI-driven automation before S/4HANA go-live

To clear a five-figure material master backlog in weeks, you must replace manual, rule-based cleaning with AI-driven semantic deduplication and automated attribute extraction. By leveraging a specialized AI engine that combines deep learning with knowledge engineering, you can categorize thousands of unstructured records and identify hidden duplicates across multiple languages in real time. This automated approach ensures a clean SAP material master data-cleaning process, allowing you to migrate only high-quality, validated data to S/4HANA.

Why material master backlogs stall S/4HANA migrations

The success of an S/4HANA migration depends entirely on the quality of the data being moved. Most enterprises discover too late that their legacy ECC system contains hundreds of thousands of material records characterized by unstructured free-text, missing attributes, and redundant entries.

In contrast to business partner data (customers or vendors), which follows a relatively predictable structure, material master data is notoriously inconsistent. For

example, a single mechanical part might be identified as 'Bearing 6205-C3 SKF' at one site and 'Ball Bearing 25mm' at another. Standard SAP Master Data Governance (MDG) serves as a robust workflow orchestrator, yet it typically necessitates manual rule creation for every category. When managing thousands of categories and millions of records, defining these rules by hand can take months or years, creating a bottleneck that jeopardizes your go-live schedule.

To achieve a "clean core," you need to solve the intelligence gap: the ability to understand what a material actually is, regardless of how it was typed or what language was used.

Step 1: Perform semantic deduplication to identify "ghost" inventory

The first step in clearing a backlog is identifying how much of it shouldn't exist in the first place. Traditional deduplication relies on "lexical matching"—looking for identical strings of text. This fails when descriptions are inconsistent.

Semantic deduplication uses AI to understand the intent and technical

specifications behind the text. By creating a semantically connected “digital twin” of your procurement data, you can identify that two different descriptions actually refer to the same physical object.

- **The benefit:** You can immediately collapse your backlog by 10–20% by identifying duplicates.
- **The supply chain impact:** Identifying these duplicates allows supply chain directors to reduce inventory value and free up working capital before the migration even begins.

Step 2: Automate categorization and technical attribute extraction

Once duplicates are identified, the remaining records must be categorized and enriched with technical attributes (such as voltage, dimensions, or material type). Doing this manually for 50,000 records is an impossible task for a data management team facing a tight deadline.

AI-driven tools, such as Creatives TAM for Master Data Governance, use proprietary knowledge engineering to automate this. Instead of a human reading a description and looking up a taxonomy code, the AI:

1. Suggests the category based on industrial standards or your internal taxonomy.
2. Extracts technical attributes directly from the free-text description.
3. Normalizes units of measure (e.g., converting “in.” to “mm” automatically).

This process transforms a manual 15-minute task per record into a sub-second automated check, allowing you to process thousands of backlog items per day.

Step 3: Implement a “clean-core” governance workflow with TAM and SAP MDG

Clearing the backlog is only half the battle; you must also ensure that “bad” data never enters the system again. The most effective architecture for this is a “Better Together” approach where Creatives TAM provides the

intelligence and SAP MDG provides the governance.

In this workflow, the AI engine acts as the “requester portal.” When a user attempts to create a new material:

- **Real-time AI check:** The AI checks for duplicates and suggests the correct category as the user types.
- **Data enrichment:** Technical attributes are extracted and validated before the record ever hits the ERP.
- **Seamless integration:** The cleaned, enriched data is sent to SAP MDG as a standard change request.

This allows SAP MDG to focus on what it does best—managing accounting, MRP, and sales views—while the AI ensures the core material data is perfect from the start.

How AI handles multilingual data backlogs

For multinational corporations, the material master backlog is often compounded by language barriers. A spare part in a German plant is described in German, while the same part in a Brazilian plant is described in Portuguese.

A global SAP material master data cleaning strategy must be language-agnostic. AI models trained on industrial domains can recognize technical specifications regardless of the language. For example, the AI knows that “Válvula” and “Valve” are the same noun and that “10 bar” is a pressure attribute in any language. This capability allows you to harmonize your entire global database into a single, searchable catalog, providing CPOs with the visibility needed for strategic sourcing.

Sustaining data quality: from backlog to business value

Clearing a five-figure backlog is a massive achievement, but the ultimate goal of S/4HANA migration is digital transformation. High-quality data is the foundation for:

- **Predictive maintenance:** Accurate material data ensures the right spare part is available when a machine fails.
- **Spend analytics:** Harmonized categories allow for precise “who bought what from whom” analysis, eliminating maverick spend.
- **Supply chain resilience:** A clean material master allows you to identify alternative suppliers and parts quickly during global disruptions.

By deploying an AI-driven solution during your migration prep, you aren’t just fixing a legacy problem—you are building a trusted data foundation for your enterprise’s future.

Accelerate your S/4HANA journey with Creactives

Creactives provides the AI intelligence that SAP MDG needs to master the material domain. By combining deep learning with a massive proprietary knowledge base of industrial components, Creactives enables you to cleanse, harmonize, and govern your data across any language and alphabet.

Whether you are struggling with a massive pre-migration backlog or looking to automate your material governance for a clean-core S/4HANA environment, Creactives offers the tools to move from data chaos to a semantically connected digital twin.

Are you ready to stop “migrating the mess”? Discover how Creactives TAM enriches SAP MDG to deliver clean, actionable material data from day one of your S/4HANA go-live:
<https://www.creactives.com/contact-us/>

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