

From 63% to 96% top-3 accuracy: How Canals built hybrid product search on Elasticsearch

Canals, which builds AI for wholesale distributors, runs a hybrid keyword-and-vector product-matching engine on Elasticsearch that lifted top-3 match accuracy from 63% to 96%, with roughly 90% of matches now auto-selected.

63% → 96%

Top-3 match accuracy, launch to today

90%

Matches auto-selected, no user action

2.8x

Turtle's quote win-rate gain on Canals

90M

Products indexed for hybrid search

The challenge

- ✗ Match a vague request, a partial part number, or a loose description to one SKU in catalogs up to 4 million items.
- ✗ Catalog knowledge lived with reps who spent years learning it, and it does not scale as catalogs grow.
- ✗ Keyword-only search hit a relevance ceiling below what the product needed.
- ✗ A wrong match means the wrong part ships: a return to process and often a lost order.

The solution

- ✓ Hybrid retrieval runs BM25 keyword, dense, and sparse vectors in one Elasticsearch index.
- ✓ Custom analyzers normalize messy part numbers and product data so variants match.
- ✓ A downstream reranking layer narrows a wide candidate net to the top matches.
- ✓ Managed Elastic Cloud runs it all, with BBQ compressing vector storage.

How it works

1. Sync & index

ERP catalog cleaned, embedded, indexed

2. Query per line item

Email, PDF, or spreadsheet parsed

3. Retrieve

BM25, dense, sparse from one index

4. Rerank

Order by best fit, apply live stock

5. Approve & return

Rep confirms; order flows to ERP

Before

- Reps relied on memory and tribal knowledge to find the right product.
- A keyword-only prototype sat below launch-quality accuracy.
- Matching at scale would need a separate vector database beside the keyword engine.

After

- The right product lands in the top three 96% of the time; about 90% auto-selected.
- Keyword, dense, and sparse candidates come from one index, narrowed by reranking.
- One managed platform; the team focuses on match quality, not the cluster.

"Elasticsearch gives you a search engine that works extremely well **out of the box**, but also gives you **all the knobs you need to tune it** for your specific domain."

— Matias Weber, Director of Machine Learning, Canals

Why it matters beyond one catalog

Next, Canals is using search to pull requirements out of drawings and spec documents and connect them to each line item, giving its AI agents a richer picture of what a customer needs. The same principles scale from 10,000 items to tens of millions: Keep keyword and vectors on one platform, cast a wide net for recall, and let a ranking layer earn the precision.