

KNOWLEDGE GRAPHS FOR E-COMMERCE

Transforming E-Commerce Search and Analytics

A Knowledge Graph Approach to Intelligent Product Discovery

A white paper by Matthew de La Fe, Senior Manager of AI, an approach for e-commerce and digital commerce leaders.

Executive Summary

E-commerce search continues to underperform expectations, leading to lower sale conversions and depressing revenues.¹ Maintaining a catalog of changing products to meet ever-changing demands and marketplace trends becomes a costly procedure each cycle. Despite adoption of AI-assisted product tagging, search continues to consistently return zero or inaccurate results, and customers regularly encounter irrelevant or incomplete search outcomes.²

These failures are not primarily caused by poor models or insufficient data. They are architectural.

Most modern AI tagging systems process products independently, extracting attributes from images and descriptions without a shared semantic foundation. As a result, two functionally identical products may be described and tagged differently depending on marketing language, image emphasis, or data availability.

This white paper presents an alternative architecture: **knowledge graph driven product discovery**. Rather than managing thousands of products as isolated records, a knowledge graph models the catalog as interconnected entities of products, materials, occasions, styles, and properties, which are governed by explicit business relationships and rules. When a product is connected to a material, it can inherit relevant properties through controlled semantic rules rather than ad-hoc tagging.

Knowledge graphs require upfront ontology design, ongoing governance, and do not eliminate every search failure. They are not a replacement for lexical or vector search engines, but a **semantic augmentation layer** that improves recall, consistency, explainability, and long-term maintenance economics.

Organizations that adopt this architecture reduce unproductive zero-result searches, significantly improve search consistency across large catalogs, and lower merchandising overhead as attributes shift from product-level tagging to ontology-level management.³

1. Current State: AI-Assisted Tagging Limitations

1.1 How AI-Assisted Tagging Works Today

Many e-commerce platforms now rely on AI to automate product tagging. Product images and descriptions are processed by vision models and large language models. Attributes such as color, style, and material are extracted. Tags are generated and applied to product records. AI-assisted tagging has reduced manual labor and improved baseline consistency for visible attributes such as color, pattern, and style.

1.2 Structural Limitations: Isolated Processing

Despite efficiency gains, AI-assisted tagging remains constrained. Each product is processed independently, without reference to a shared domain knowledge. As a result, attributes depend heavily on product descriptions. Semantically equivalent products may surface for different queries. Search relevance can vary across the catalog.

Two linen shirts may be tagged differently depending on whether the description emphasizes “summer wear” or “relaxed fit.” It is not inferred that linen, by its nature, is commonly breathable and cooling fabric.

Product A: “Classic linen shirt perfect for warm weather”

AI Tags: linen, classic, warm weather, summer

Product B: “Relaxed fit shirt in premium linen”

AI Tags: linen, relaxed fit, premium

A customer searching “summer shirt” finds Product A but not Product B. Advanced post-processing rules and co-occurrence models can address these issues, but they effectively recreate ontology-like behavior implicitly based on training data. Synonym lists can balloon to unmanageable levels, and explanation is hidden behind probabilities. Maintaining a large data catalog becomes burdensome, especially in industries with short cycles. The conclusion is not that AI cannot reduce this gap, but that doing so at the product level is **uneconomical and ungovernable** at scale.

1.3 The Impact on Search Performance

These architectural constraints manifest directly in search. Consumers are leaving sites after negative search experiences.⁴ In an era of digital dominance, e-commerce players must treat the online experience with as much reverence as in-store experiences of the past. A poorly architected search is the equivalent unorganized, unkempt displays.

2. The Knowledge Graph Alternative

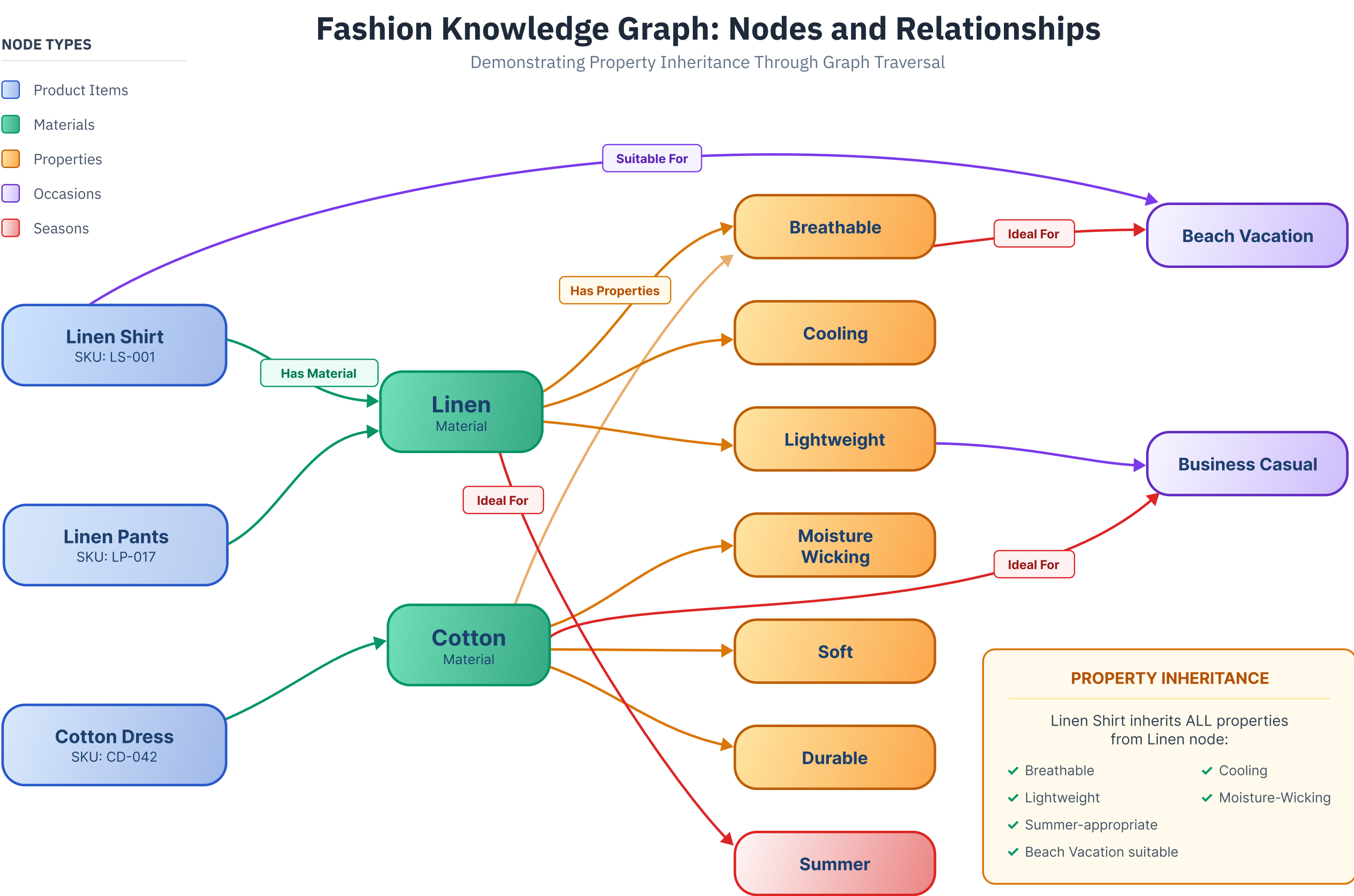
Knowledge graphs offer a fundamentally different architecture. Major retailers such as, Amazon, Walmart, Ikea, and Zalando have adopted a graph-based approach for product catalogs.

2.1 Core Concepts: Entities, Properties, and Relationships

A knowledge graph represents catalog knowledge as a network of entities governed by an ontology:

- Entities (Nodes):** Products, materials, categories, occasions, climates, styles
- Properties (Node Attributes):** Attributes attached to entities (e.g., breathable, lightweight, formal)
- Relationships (Edges):** Explicit links between entities (e.g., hasMaterial, suitableFor, idealFor)

This structure allows knowledge to be defined once and reused consistently across the catalog.



2.2 Controlled Property Inheritance

Property inheritance is the key differentiator. When a product is linked to a material such as linen, relevant properties may be inherited subject to rules, such as:

- Fabric Weights:** Heavy linen may not qualify as “lightweight”
- Construction:** Linen + lining ≠ breathable; linen blazer ≠ summer shirt
- Blends:** A 20% linen blend does not inherit full linen properties
- Category:** Same material may have different property applicability by garment type

Properties are inherited based on material relationships, subject to ontology rules and exclusions. Constraint validation frameworks prevent inappropriate propagation. This *rule-governed* inheritance ensures semantic accuracy while preserving consistency.

2.3 Maintenance Economics

Traditional tagging scales linearly with catalog size. More importantly, when customer language changes (for example, “eco-friendly” becomes a dominant search), updates occur once at the ontology level and propagate systematically without touching product records.

3. The Role of AI in a Knowledge Graph Architecture

Knowledge graphs do not eliminate AI; they reposition it.

Instead of asking AI to generate open-ended tags per product, AI is deployed for constrained tasks that operate at the semantic level.

3.1 Entity Classification (Constrained AI)

Instead of AI-assisted tagging, AI classifies products against known ontology entities. Predictions are confidence-scored, routed through **human-in-the-loop review workflows**, and continuously evaluated. This is a fundamentally easier AI task with higher accuracy because the output space is constrained to known entities with clear definitions.

Traditional AI: “What tags describe this product?”

Output: Unlimited possibilities, inconsistent across products, no governance

Knowledge Graph AI: “Classify against our ontology:”

Material: Linen (94%) → review queue → approved → all linen properties inherited

Output: Constrained, verifiable, auditable, governed

3.2 Ontology Enrichment

Our approach includes analyzing search logs, product descriptions, and external content to propose new properties, emerging terminology, and gaps in semantic coverage. Behavior data suggests new relationships between nodes.

AI ONTOLOGY ANALYSIS EXAMPLE

- Detection:** "moisture-wicking shirts" has 10,000 monthly searches with no ontology match
- Analysis:** 847 products reference this in descriptions; linked to Polyester (412), Merino (203), Nylon (232)
- Recommendation:** Add property "moisture-wicking" to 3 material nodes
- Governance:** Review and approve. Property propagates to 847 products.

3.3 Query-to-Ontology Mapping

Our approach maps natural language queries to ontology concepts, enabling disambiguation, multi-concept resolution, and explainable query interpretation. This mapping layer improves continuously through machine learning on click-through data and query refinement patterns.

4. Knowledge Graph Impact and Implementation

4.1 Zero-Result Reduction

Knowledge graphs reduce zero-result searches by creating multiple semantic paths to products. ⁵

4.2 Hybrid Search Enablement

In production, knowledge graphs augment existing infrastructure

- Lexical search (BM25, inverted indexes):** Handles keyword matching, typo tolerance, stemming
- Vector retrieval (embeddings):** Captures semantic similarity for long-tail queries
- Knowledge graph reasoning:** Provides governed semantic expansion and property inheritance
- Ranking, personalization, merchandising rules:** Business logic preserved exactly as-is

Knowledge graphs complement lexical and vector retrieval layers. The graph provides a semantic backbone to deliver auditable, controlled, and business-aligned interpretation.

4.3 Where Knowledge Graphs Can Fail

- Poor Design:** Ontologies must reflect the true nature of the consumer search
- Missing query mapping layer:** Queries must map to ontology concepts
- Over-connection:** Trade-offs between precision and recall must be made
- Novel Concepts:** Unique search trends require ontology manipulation
- Ontology errors:** Governance structure is essential for maintenance

Zero-Result Cause	KG Impact	Improvement
Synonym Gaps	Property Inheritance	High
Incomplete Tagging	Relationship Hopping	High
Novel Terminology	AI maps to existing concepts	Medium-High
Misspellings	Requires lexical fuzzy matching	None (same)
Trend-driven Noise	Ontology enrichment (delayed)	Medium

5. Business Impact

5.1 Operational Efficiency

- Faster product onboarding:** Classify against ontology vs. comprehensive tagging
- Faster trend response:** Update ontology structure (one change) vs. retag products (thousands of changes)
- Consistency guaranteed by structure:** Not dependent on individual tagging accuracy or AI-assisted training
- Improved auditability and governance:** All semantic decisions are traceable and reviewable

6. Implementation Considerations

- Ontology design requires cross-functional input from search, merchandising, data science, and domain experts. Governance ownership must be explicit.
- Measurement should track relevance and meaningful query resolution, not just zero-result counts. Continuous refinement is expected.

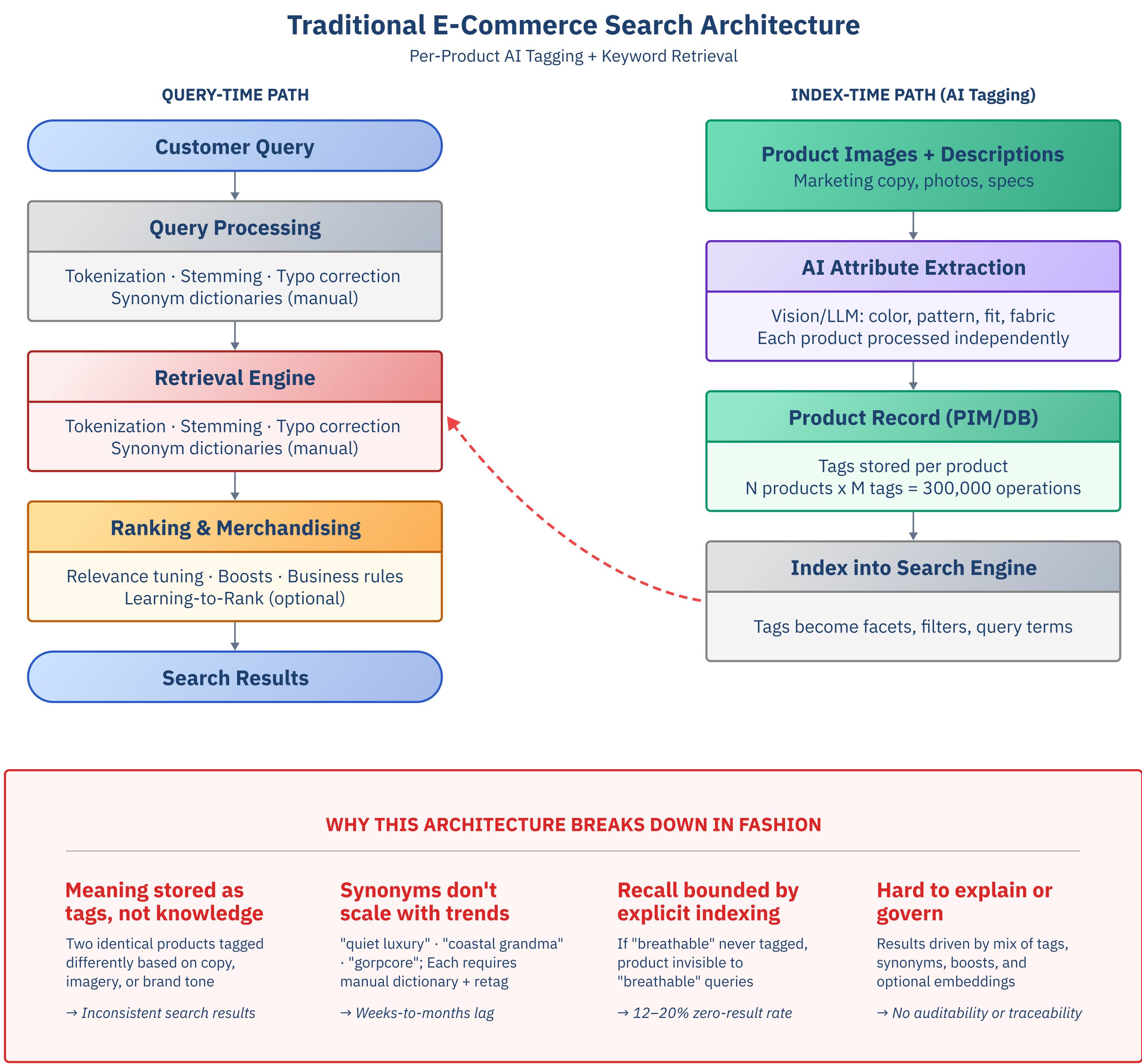
6.1 Governance and Ownership

- A knowledge graph requires governance.
- Ontology Stewards:** Approve structural changes to the graph
- Workflow Reviewers:** Approve AI proposed propagation
- Management Trackers:** Measure KPIs against traditional search conversion

7. Architecture: Traditional vs. Knowledge Graph

The following diagrams compare the traditional e-commerce search architecture with the knowledge graph–augmented approach. The knowledge graph does not replace existing infrastructure. It augments it by providing a governed semantic layer.

7.1 Traditional E-Commerce Search Architecture



7.2 Knowledge Graph Augmented Architecture

The knowledge graph–augmented stack preserves all existing infrastructure while adding a semantic layer that improves recall, consistency, and explainability.

7.3 Architectural Benefits

- Recall without noise:** Increases findability and preserve precision ⁶
- Explainability:** Results can be traced to explicit semantic relationships
- Resilience:** Customer language evolves without requiring product re-tagging
- Composable:** Works with existing commerce platforms and search engines

Dimension	Traditional	KG-Augmented
Retrieval depends on	Per-product tags + dictionaries	Governed entities + relationships
Failure mode	One missed tag = missed product	Ontology gap = fix once, propagate
Maintenance	N products × M tags	X entities × M properties + N classifications
Explainability	Opaque score blending	Traceable semantic relationships
Trend response	Manual retag + synonyms	Ontology update, auto-propagation
AI role	Open-ended tag generation	Constrained classification + enrichment

8. Addressing Common Objections

Objection 1: *"Isn't this just moving the tagging problem somewhere else?"*

Response: The problem is not tagging itself, but where tagging occurs. Product-level tagging scales linearly with catalog size and is inherently inconsistent. Ontology-level definition scales with domain complexity, not SKU count, and is governed by explicit rules and review workflows. One ontology update to a material node propagates to every product using that material without touching product records.

Objection 2: *"Won't this over-generalize products and hurt precision?"*

Response: Only if implemented incorrectly. Property inheritance is rule-governed, contextual, and category-aware. Constraint validation prevents inappropriate propagation (e.g., a lined linen blazer does not inherit "summer-only" attributes). Precision is preserved through validation rules and downstream ranking layers.

Objection 3: *"We already use embeddings. Don't they solve this?"*

Response: Embeddings improve recall but lack explicit semantics, governance, and explainability. A vector search may surface a linen shirt for "breathable," but it cannot explain why, cannot be audited, and cannot be governed by business rules. Knowledge graphs provide the semantic backbone that embeddings alone cannot: auditable meaning, controlled inheritance, and business-aligned interpretation. They are complementary, not competitive.

Objection 4: *"This sounds expensive to maintain."*

Response: The opposite is typically true. Maintenance shifts from thousands of individual product records to a small, governed set of entities and rules. Over time, operational effort decreases as ontology changes propagate automatically. ⁷

Objection 5: "Our catalog and brands are too unique for a single ontology."

Response: Ontologies are extensible by design. Brand-specific semantics, exclusions, and overrides are expected and supported. A shared core ontology enables cross-catalog consistency, while extensions preserve differentiation.

Objection 6: "What happens when the ontology is wrong?"

Response: Ontology errors are visible, traceable, and correctable in one place. Compare this to traditional tagging, where silent errors are distributed across thousands of SKUs with no auditability. A governed ontology with stewardship workflows is fundamentally safer than unmonitored product-level tagging at scale.

Conclusion

AI-assisted tagging has reached diminishing returns in e-commerce search. The limitation is not model quality, but architecture.

Knowledge graphs provide a governed semantic foundation that enables AI to operate at the structural level to improve consistency, explainability, and long-term economics. When implemented as part of a hybrid search stack, they deliver meaningful reductions in search failure while positioning organizations for continuous improvement as customer language evolves.

The honest assessment: this is not a magic solution. It requires upfront investment in ontology design, ongoing governance, and clear organizational ownership. But it fundamentally changes the economics of catalog management and creates a resilient, auditable architecture that scales with domain complexity rather than SKU count.

Next Steps

Cyberhill Partners specializes in ontology-driven e-commerce architectures. We can assess your current search architecture, quantify semantic gaps, design a custom ontology and implement AI-powered governance systems.

Contact us for a complimentary e-commerce search and semantic architecture assessment.

[Schedule an AI E-commerce Strategy Session](#)

[Learn more at cyberhillpartners.com](https://cyberhillpartners.com)

Appendix: Technical Reference

A.1 Sample SPARQL Query

```
SELECT ?item ?itemName ?material
WHERE {
  ?item a :Product ;
        :hasName ?itemName ;
        :hasMaterial ?material ;
        :suitableFor :BeachVacation .
  ?material :hasProperty :Breathable .
}
```

A.2 AI Roles Summary

AI Role	Function	Governance
Entity Classification	Constrained classification against ontology	Review workflows, brand-specific overrides
Ontology Enrichment	Discover new properties from search logs	Steward approval before propagation
Relationship Discovery	Identify connections from behavioral data	Advisory only; domain expert approval
Query Mapping	Map natural language to ontology concepts	Click-through learning, continuous monitoring

A.3 Glossary

- Knowledge Graph:** A structured representation of entities, properties, and relationships stored as triples
- Ontology:** A formal specification of entity types, relationship types, property definitions, and constraint rules
- Triple:** Subject → Predicate → Object (e.g., Linen_Shirt → hasMaterial → Linen)
- Node:** An entity in the graph (product, material, occasion, property)
- Edge:** A relationship connecting two nodes (hasMaterial, hasProperty, idealFor)
- Property Inheritance:** Rule-governed transfer of properties through relationships, subject to constraints
- Graph Traversal:** Navigating connected nodes to expand recall beyond keyword matching
- Ontology Steward:** Named individual or team responsible for reviewing and approving ontology changes
- Constraint Validation:** Rules that prevent inappropriate property propagation (e.g., SHACL)

References

¹ Tharp, C. (2023, March 23). New research: Search abandonment continues to vex retailers worldwide. Google Cloud. <https://cloud.google.com/blog/topics/retail/new-research-on-search-abandonment-in-retail>

¹ Holst, C. (2014, August). The current state of e-commerce search. Smashing Magazine. <https://www.smashingmagazine.com/2014/08/the-current-state-of-e-commerce-search/>

² Rouse, E. (2023, February). The future of ecommerce search. <https://www.nosto.com/blog/new-search-research/>

Germain, J. (2025, April). Why Smart Search Matters: The Key to Higher E-Commerce Conversions. <https://www.ecommercetimes.com/story/why-smart-search-matters-the-key-to-higher-e-commerce-conversions-178245.html>

³ Dong, X. L. (2018). Challenges and innovations in building a product knowledge graph. <https://dl.acm.org/doi/10.1145/3219819.3219938>

³ Dong, X. L. (2020). Build Product Graphs Automatically. <https://www.amazon.science/blog/building-product-graphs-automatically>

³ Cheng, B., et al. (2023, March 21). How CSC Generation powers product discovery with knowledge graphs using Amazon Neptune. <https://aws.amazon.com/blogs/database/how-csc-generation-powers-product-discovery-with-knowledge-graphs-using-amazon-neptune/>

³ Deivasigamani, K. (2020, March 13). Retail Graph—Walmart’s product knowledge graph. <https://medium.com/walmartglobaltech/retail-graph-walmarts-product-knowledge-graph-6ef7357963bc>

⁴ Tharp, C. (2023, March 23). New research: Search abandonment continues to vex retailers worldwide. Google Cloud. <https://cloud.google.com/blog/topics/retail/new-research-on-search-abandonment-in-retail>

⁴ Rouse, E. (2023, February). The future of ecommerce search. <https://www.nosto.com/blog/new-search-research/>

⁴ Retail Customer Experience. (2021 September 30). Challenging product search top reason consumers bounce off e-commerce sites. <https://www.retailcustomerexperience.com/news/challenging-product-search-top-reason-consumers-bounce-off-e-commerce-sites/>

⁴ Storyblok. (2022, December 20). 60% of consumers abandon purchases due to poor website user experience, costing e-commerce companies billions. <https://www.storyblok.com/mp/poor-website-user-experience>

⁵ Sarika, J. (2020, October 11). Exploiting knowledge graphs for facilitating product/service discovery. <https://arxiv.org/abs/2010.05213>

⁵ Yu, C., et al. (2023, July). FolkScope: Intention knowledge graph construction for e-commerce commonsense discovery. <https://aclanthology.org/2023.findings-acl.76/>

⁵ Yu, C., et al. (2024, May 10). Building commonsense knowledge graphs to aid product recommendation. <https://www.amazon.science/blog/building-commonsense-knowledge-graphs-to-aid-product-recommendation>

⁵ Xv, G., et al. (2023, August 4). E-commerce search via content collaborative graph neural network. <https://dl.acm.org/doi/10.1145/3580305.3599320>

⁶ Zalmout, N., et al. (2021). All you need to know to build a product knowledge graph. https://naixlee.github.io/Product_Knowledge_Graph_Tutorial_KDD2021/

⁶ Xu, D., et al. (2020, January 22). Product knowledge graph embedding for e-commerce. <https://dl.acm.org/doi/10.1145/3336191.3371778>

⁶ Dong, X. L. (2018). Challenges and innovations in building a product knowledge graph. <https://dl.acm.org/doi/10.1145/3219819.3219938>

⁷ Moosavi, S., Nematbakhsh, M., & Farsani, H. K. (2012). Designing a catalog management system—An ontology approach. Expert Systems with Applications. <https://www.semanticscholar.org/paper/a4487a2379172ac2db6159f7fa87501d14eafc88>

Cheng, B., et al. (2023, March 21). How CSC Generation powers product discovery with knowledge graphs using Amazon Neptune. <https://aws.amazon.com/blogs/database/how-csc-generation-powers-product-discovery-with-knowledge-graphs-using-amazon-neptune/>

da Cruz, T., et al. (2025, November 8). Ontology Learning and Knowledge graph Construction: A Comparison of Approaches and Their Impact on RAG Performance. <https://arxiv.org/html/2511.05991v1>

Angermann, H., & Ramzan, N. (2024). Ontologies and reasoning to capture product complexity in automation industry. https://www.academia.edu/84527745/Ontologies_and_Reasoning_to_Capture_Product_Complexity_in_Automation_Industry