

Leveraging Large Language Models and Machine Learning to Tackle Challenges in E-Commerce Catalog Management and Its Implications on Advertising

¹Anoop Singh Saini, ²Mahesh Chayel Mohinder Singh

Submitted: 20/02/2024 Revised: 25/03/2024 Accepted: 08/04/2024

Abstract: This paper highlights the challenges e-commerce platforms face in managing large catalogs and how Large Language Models (LLMs) can be leveraged to drive operational efficiency, customer engagement, and personalized digital Ads. By analyzing current research and industry practices, we showcase that well-structured catalogs can significantly reduce cart abandonment rates and improve customer satisfaction, since the vast majority of shoppers consider product content crucial in purchasing decisions. The paper details how LLMs and Machine Learning Models can transform traditional catalog management by automating product classification, optimizing attribute extraction and content accuracy, while enhancing search accuracy. We also present a framework for leveraging enhanced catalog data in advertising through three primary techniques: personalized product recommendations, focused retargeting campaigns, and concurrent dynamic pricing. Our findings suggest that businesses implementing LLM and ML-driven catalog management solutions can achieve superior customer engagement and operational efficiency. Our objective with this paper is to contribute to knowledge on catalog management and suggest some scalable catalog management strategies.

Key words: E-commerce, Product Catalog, Digital Advertising, GenAI, Artificial Intelligence, LLMs

Introduction:

In the context of e-commerce, a “catalog” refers to a comprehensive collection of products or services that a business offers online. This digital catalog serves as a virtual storefront, allowing customers to browse, search, and select items for purchase. Unlike traditional catalogs, which may be printed and static, online catalogs are dynamic and interactive, providing real-time information such as prices, availability, and detailed product descriptions. They are essential for facilitating the online shopping experience, as they not only showcase products but also help customers make informed purchasing decisions by presenting all relevant information in an organized manner (Catsy, n.d.), (DCatalog, n.d.).

Effective catalog management is a fundamental aspect of e-commerce that significantly influences customer experience and business performance. In the digital marketplace, catalogs serve as the primary data source through which customers discover and interact with products online or via mobile apps. They encompass detailed product information, including descriptions, product characteristics, images, and other assets, which are essential for helping consumers make informed purchasing decisions. The organization and maintenance of this information are critical; a well-structured catalog not only enhances user experience but also drives sales, customer satisfaction, operational efficiency, and regulatory compliance.

1Independent Researcher, sainianoop@gmail.com

2Independent Researcher, mahesh.chayel@sloan.mit.edu

Catalog plays one of the key roles in advertising products, and a poorly set up catalog can result in poor ad relevance, lower quality scores, lower personalization, and inefficient ad spending.

A Brief Review of Literature: Importance of Catalog Management

The management of an e-commerce catalog is critical for ensuring that product data is accurate, up-to-date, and consistently presented across various sales channels. This process involves organizing product listings, categorizing items appropriately, and maintaining high-quality images and product characteristics/attributes. Effective, comprehensive, and accurate product catalogs enhance site experience capabilities like item page, search, and personalization making it easier for customers to find what they are looking for quickly and easily. Research indicates that a well-structured catalog can significantly reduce cart abandonment rates and improve overall customer satisfaction (Clarity Ventures, n.d.), (ShipBob, n.d.). The relationship between a well-structured catalog and reduced cart abandonment rates is rooted in several key factors that enhance the overall shopping experience. Some of the ways through which this takes place are:

- Simplified Navigation and Accurate Information

A well-structured catalog ensures that product information is accurate, up-to-date, and easily accessible. This clarity helps customers navigate the website efficiently, reducing the likelihood of frustration and abandonment. Reducing the time and effort required by customers to locate desired products/items has a direct impact on customer

experience, resulting in improved conversion rates. For example, surfacing all the similar/variants of a product on a single page can make it easier for customers to discover items and ease the purchasing decision-making. A research conducted by the Baymard Institute suggests that a simplified checkout process, such as offering a single-page option, can reduce cart abandonment rates by up to 35% (ContactPigeon, n.d.).

- **Minimizing Decision Fatigue**

When e-commerce sites have millions of items to choose from, it becomes overwhelming for customers, which further results in a sub-optimal customer experience. In their study, Barilliance highlights that excessive choices often lead to indecision. As mentioned above, grouping similar products together and offering useful filters, companies can enable customers to effectively discover and purchase products (Barilliance, n.d.).

- **Fostering Customer Confidence**
Complete and accurate product information, along with transparency on pricing and shipping fees, are some of the common factors that determine whether the customer would feel confident in making the purchasing decision or not. Research also shows that by addressing these concerns, businesses can ease customer anxieties and boost conversion rates (ContactPigeon, n.d.) (ViSenze, 2023).

- **Optimizing for Mobile Users**

With mobile devices driving a large share of e-commerce traffic, it's essential for catalogs to be mobile-friendly. Mobile shoppers have the highest rates of cart abandonment, with over 85% of transactions failing to convert into sales. This is often due to challenges like smaller screens and poor usability. By designing catalogs that work seamlessly on mobile devices, businesses can enhance the shopping experience and lower abandonment rates. According to Hotjar, mobile devices account for more than two-thirds of e-commerce traffic, underscoring the importance of mobile optimization in reducing cart abandonment.

Adapting to Inventory Changes

A streamlined catalog enables businesses to quickly adapt to shifts in inventory or product offerings, ensuring customers always have access to accurate information (Vue.ai, 2023). Effective catalog management is vital for maintaining consistency and responsiveness in e-commerce operations. Here are some strategies businesses use to achieve this:

- **Real-Time Inventory Tracking**

One of the key advantages of a well-maintained catalog is the ability to update inventory in real time. This ensures customers see accurate stock levels, avoiding issues like overselling or displaying unavailable items. Inventory management tools are essential for this process. For example, platforms like ShipBob provide real-time inventory data, allowing businesses to monitor stock levels and receive alerts when supplies run low, helping to prevent stockouts (Lopienski, 2023).

- **Streamlined Data Updates**

Automated catalog systems simplify the process of updating product details, such as pricing, descriptions, and specifications. These systems can perform data quality checks automatically, ensuring consistency across all sales channels. Automation reduces the need for manual updates and minimizes errors. BigCommerce notes that catalog management software centralizes product data, runs quality checks, and optimizes listings for search engines, helping businesses maintain accurate and current product information (Guide, n.d.).

- **Unified Inventory Management**

Integrating the catalog with an Enterprise Resource Planning (ERP) system is crucial for aligning physical and digital inventory. An ERP system synchronizes stock levels, pricing, and product details, ensuring that online listings match actual inventory. This integration allows businesses to respond swiftly to changes, improving the overall customer experience. As CommerceBuild explains, combining physical and digital catalogs through an ERP system eliminates discrepancies, streamlines operations, and provides real-time insights into product performance (CommerceBuild, n.d.).

- **Regular Audits and Updates**

Regular inventory audits are a fundamental part of maintaining an accurate and up-to-date catalog. These audits verify the accuracy of product information, compare data to actual inventory, and update the catalog as necessary. By conducting these audits, businesses can identify discrepancies, update product information, and ensure that the catalog reflects the current state of their inventory. Shopify recommends maintaining a single source of truth for all product information and regularly auditing catalogs to ensure data accuracy and consistency across multiple sales channels (Guide, 2024).

Finally, this paper implements the thematic findings with the help of a use case.

Results and Discussion: Challenges in Catalog Management

Despite its critical importance, e-commerce catalog management presents several challenges that can hinder operational efficiency and customer satisfaction. As businesses scale and diversify their offerings, keeping product catalogs organized and consistent across various platforms becomes increasingly difficult. Expanding product lines lead to a surge in variations, categories, and metadata, which, if not properly managed, can result in disorganization. A report by BigCommerce highlights that as online retailers grow, they frequently struggle with maintaining standardized product information—an essential factor for ensuring accuracy and consistency across different sales channels (guide, n.d.).

Managing a growing catalog presents several challenges, especially when selling across multiple platforms. Each marketplace—whether proprietary or third-party, such as Walmart, Amazon, or eBay—has distinct categorization and formatting requirements. Without proper standardization, businesses risk listing products incorrectly, leading to misclassification, incomplete descriptions, and inconsistencies that can frustrate customers. Research by BigCommerce indicates that without structured metadata, product listings may suffer from inaccuracies, negatively impacting search visibility and the overall shopping experience (TransForm Solutions, n.d.).

Updating and maintaining a catalog manually is a time-consuming process, requiring constant revisions to reflect inventory changes, pricing adjustments, and new product introductions. Many retailers struggle to keep product information well-structured due to the vast volume of data they must manage. Studies also point to businesses not investing enough in keeping their specs updated, ensuring data migration from one version of spec to another, and appropriate categorization of items, which results in incomplete or misleading product listings (guide, n.d.). These inefficiencies negatively contribute to broader customer experience and businesses's ability to win customer's trust.

Leveraging LLMs and Machine Learning For Catalog Optimization

To mitigate these challenges, many businesses are turning to Machine Learning (ML) to automate catalog management, reducing manual effort and improving accuracy. ML-powered solutions enhance the organization of product information, ensuring seamless integration across multiple platforms. According to Netscribes, AI-driven catalog management automates data processing,

reducing inconsistencies and minimizing human errors (Netscribes, 2023).

The benefits of ML extend across several key areas:

- **Automated Product Classification:** ML algorithms analyze product attributes, descriptions, and images to ensure proper categorization, improving search functionality and discoverability (Anolytics.ai, n.d.).
- **Enhanced Attribute Extraction:** AI techniques help identify and fill in missing product attributes, leading to more complete and structured catalog data (Sinha & Gujral, 2024).
- **Improved Search and Personalization:** Intelligent classification improves search relevance and enables personalized product recommendations, enhancing the customer experience (Rishabh Software, 2021).

Intelligent Product Categorization

A major advantage of ML in catalog management is its ability to classify products with high accuracy. Advanced deep learning techniques and natural language processing (NLP) allow systems to analyze text and visual data to ensure proper categorization (GitHub, n.d.). This improves product discoverability, ensuring that customers can find what they need more easily. Additionally, AI-powered classification minimizes listing errors, enhancing operational efficiency and helping retailers maintain consistency across multiple platforms.

Extracting and Standardizing Product Attributes

Inconsistent product attributes create significant challenges in catalog management. The SANTA (Scalable Approach for Normalizing E-commerce Text Attributes) framework demonstrates how ML can standardize different attribute representations, such as converting terms like "HD," "1280x720," and "720p" into a unified format (Amazon Science, n.d.). Given that different sellers may describe the same product in various ways, ML helps harmonize these descriptions, reducing confusion and improving search results.

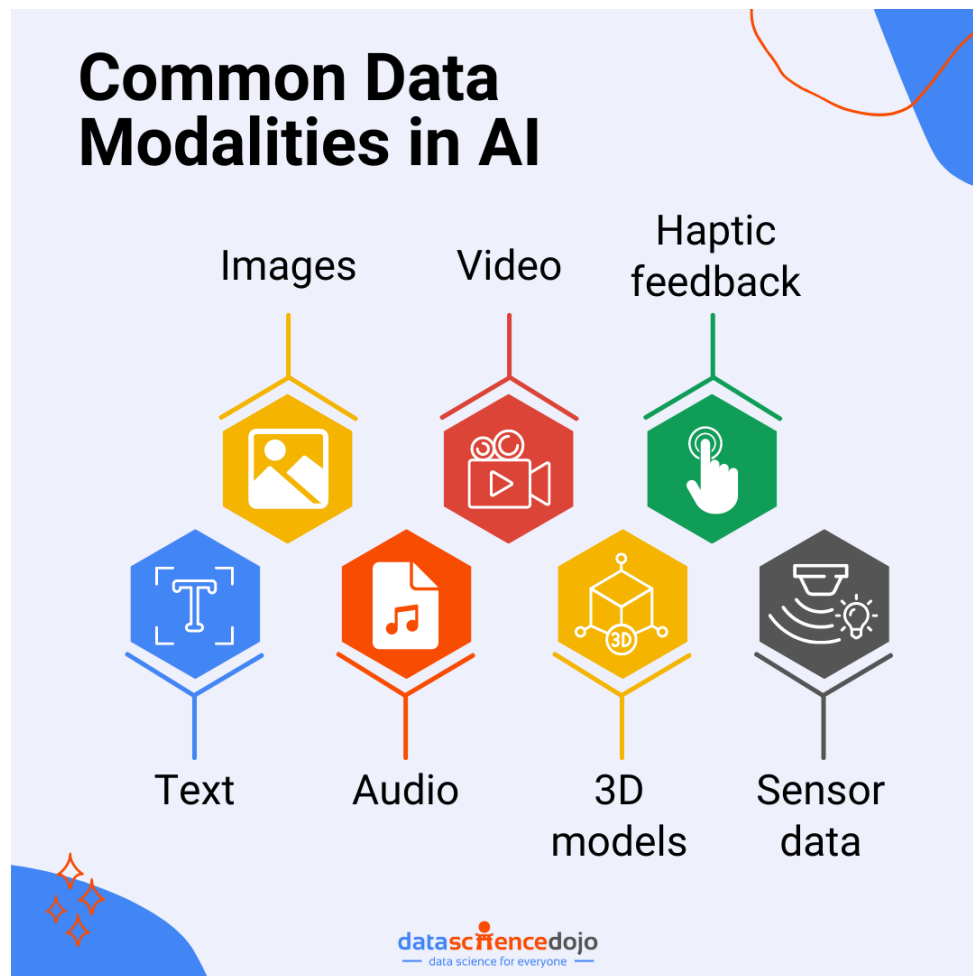
Recent advancements in attribute extraction leverage multimodal AI techniques, allowing systems to analyze textual descriptions, images, and PDFs to identify missing product details with high accuracy (Sinha & Gujral, 2024). These models employ neural networks and language processing to ensure consistency, making catalog management more efficient.

Enhancing Product Clustering and Bundling

ML also facilitates product clustering, which improves customer experience by grouping similar items together. Algorithms such as K-means

clustering enable retailers to create product bundles based on purchasing patterns. For example, e-commerce platforms can use AI to recommend complementary products, such as pairing phone

cases with screen protectors (Experro, n.d.). This enhances product discoverability and improves conversion rates by offering relevant suggestions to shoppers.



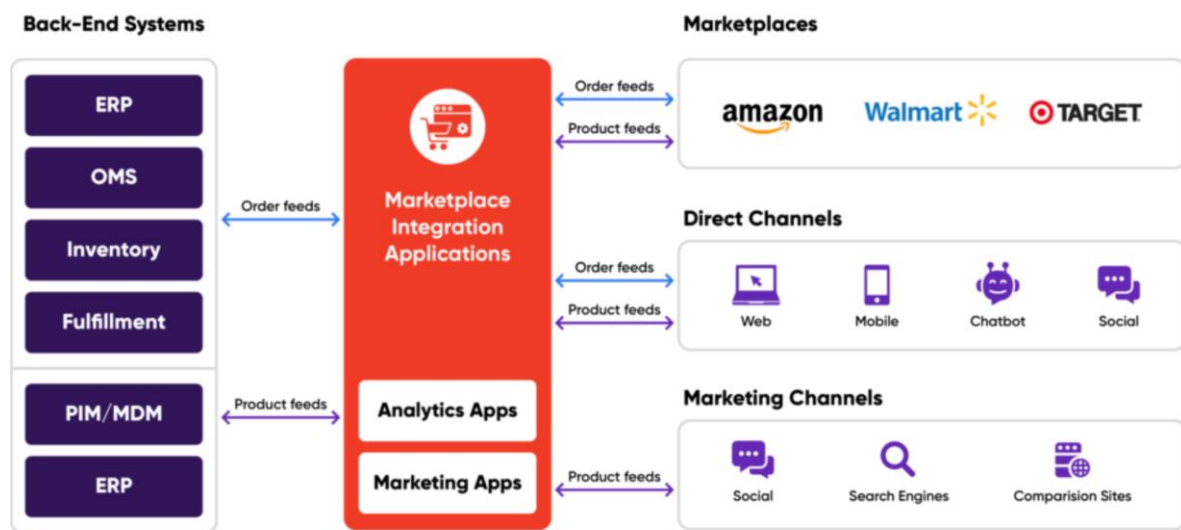
Multi-modal LLMs can be really powerful in extracting attribute values from various data sources like text, images, and video at scale. These LLMs can be used either in conjunction with ML Models to test the output, or, as LLMs become more cost-effective, to replace the hundreds of ML models. The benefits of being able to leverage LLMs, assuming the outcomes are of desired quality, can be significant in reducing the burden of creating and maintaining hundreds of ML models.

Example: Leveraging Improved Catalog For Advertising

One of the biggest benefits of better catalog management is more effective and targeted advertising. Below we recommend three ways that we can leverage the enhanced catalog.

1. Personalized product recommendations: By using the catalog's up to date information and understanding product hierarchy, it will be possible to
 - a. Bring ML to analyze customer behavior and history, and then match it with product attributes available from the catalog.
 - b. Render dynamic ads that showcase products based on the preferences of individual customers and their behaviors.
2. Focused retargeting campaigns: The retargeting strategies can be improved by:
 - a. Creating lower funnel ads that leverage the upper funnel interaction by the customers. A combination of the full funnel ads can be enhanced by using structured catalog data on top of prospecting ads.
 - b. Real time inventory data can then be used to further target based on in-stock items with price to increase ad efficiency and decrease customer frustration.
3. Concurrent dynamic pricing: Concurrent pricing information from catalog can be used for
 - a. Implementing dynamic pricing which uses features like inventory levels, competitor pricing and demand fluctuations.

- b. Using standardized product attributes to target different platforms while maintaining a consistent brand image.



Catalog management across marketplaces (Pimcore, n.d.)

Conclusion

In the dynamic landscape of e-commerce, catalog management has emerged as a critical component of business success, directly influencing customer experience, operational efficiency, and brand reputation. The evolution of digital catalogs from static, printed listings to interactive, real-time platforms demonstrates the transformative power of technology in retail. As businesses navigate increasingly complex product ecosystems, the importance of robust catalog management becomes paramount.

The study of catalog management reveals several fundamental insights:

- **Customer-Centric Approach:** 87% of shoppers consider product content crucial in purchasing decisions, underscoring the need for comprehensive and accurate product information.
- **Technological Innovation:** Machine Learning and automated systems are revolutionizing catalog management, enabling businesses to handle complex product listings with unprecedented efficiency.
- **Strategic Importance:** Well-managed catalogs can reduce the cart abandonment rates, build customer trust, and drive long-term business growth.

As e-commerce continues to evolve, catalog management will likely become even more sophisticated. Emerging technologies like AI and machine learning will play increasingly critical roles in:

- Automating product categorization
- Ensuring data consistency
- Personalizing customer experiences
- Streamlining multi-channel sales strategies

Businesses that invest in advanced catalog management solutions will be better positioned to meet the growing expectations of digital consumers, ultimately differentiating themselves in a competitive marketplace especially ads where we see three primary use cases emerge:

- Personalized product recommendations
- Focused retargeting campaigns
- Concurrent dynamic pricing

The rapid pace of GenAI improvements also mean that multi-modal LLMs can be really powerful in extracting attribute values from various data sources like text, images, audio, and video at scale. These LLMs can be used either in conjunction with ML Models to test the output, or, as LLMs become more cost-effective, to replace the hundreds of ML models. The benefits of being able to leverage LLMs, assuming the outcomes are of desired quality, can be significant in reducing the burden of creating and maintaining hundreds of ML models.

References

- [1] ., Y. M. (2024, September 22). *The Importance of Product Description: 9 Statistics & Facts*. BusinessDasher. Retrieved February 1, 2025, from <https://www.businessdasher.com/importance-of-product-description/>
- [2] ACL Anthology. (2023, July 10). *Large Scale Generative Multimodal Attribute Extraction for E-commerce Attributes*. ACL Anthology. Retrieved February 1, 2025, from <https://aclanthology.org/2023.acl-industry.29.pdf>
- [3] Amazon Science. (n.d.). *Learning Variant Product Relationships and Variation Attributes from E-commerce Website Structures*. Amazon Science. <https://assets.amazon.science/88/64/d9d6249e4bd899c5114fe869ff5d/learning-variant-product-relationship-and-variation-attributes-from-e-commerce-website-structures.pdf>
- [4] Amazon Science. (n.d.). *Scalable Approach for Normalizing E-commerce Text Attributes (SANTA)*. Amazon Science. <https://assets.amazon.science/f0/4d/03fd2f9c480f846046b9153c085a/scalable-approach-for-normalizing-ecommerce-text-attributes.pdf>
- [5] Anchanto. (n.d.). *Guide to Catalog Management System Software | E-commerce Product Catalog Management*. Anchanto. Retrieved February 1, 2025, from <https://anchanto.com/catalogue-management-101-a-complete-guide-for-your-ecommerce-business/>
- [6] Analytics.ai. (n.d.). *E-commerce Machine Learning Product Classification Insight*. Analytics.ai. <https://www.analytics.ai/blog/e-commerce-machine-learning-product-classification-insight/>
- [7] arXiv. (2024, May 27). *PAE: LLM-based Product Attribute Extraction for E-Commerce Fashion Trends*. arXiv. Retrieved February 1, 2025, from <https://arxiv.org/html/2405.17533v1>
- [8] Barilliance. (n.d.). *Cart Abandonment Rate Statistics*. Cart Abandonment Rate Statistics. <https://www.barilliance.com/cart-abandonment-rate-statistics/>
- [9] Beer, S. (n.d.). *ECommerce Product Catalog Management: Benefits & Best Practices*. Clarity Ventures. Retrieved February 1, 2025, from <https://www.clarity-ventures.com/ecommerce/ecommerce-product-catalog-management-benefits>
- [10] Bestarion. (n.d.). *Bestarion*. Applying Machine Learning in Product Matching. <https://bestarion.com/applying-machine-learning-in-product-matching/>
- [11] BigCommerce. (n.d.). *Catalog Management: What It Is and Why It's Important*. BigCommerce. Retrieved February 1, 2025, from <https://www.bigcommerce.com/articles/ecommerce/catalog-management/>
- [12] Business Dasher. (n.d.). *Importance of Product Description in Ecommerce*. Business Dasher. <https://www.businessdasher.com/importance-of-product-description/>
- [13] Catsy. (n.d.). *Ecommerce Catalog Management: The Ultimate Guide*. Ecommerce Catalog Management: The Ultimate Guide. <https://catsy.com/blog/ecommerce-catalog-management/>
- [14] Clarity Ventures. (n.d.). *Ecommerce Product Catalog Management Benefits*. Ecommerce Product Catalog Management Benefits. <https://www.clarity-ventures.com/ecommerce/ecommerce-product-catalog-management-benefits>
- [15] CommerceBuild. (n.d.). *Best Practices for Ecommerce Catalog Management*. CommerceBuild. <https://commercebuild.com/blog/best-practices-for-ecommerce-catalog-management/>
- [16] ContactPigeon. (n.d.). *Cart Abandonment Stats: The Ultimate Collection of Data and Trends*. Cart Abandonment Stats: The Ultimate Collection of Data and Trends. <https://blog.contactpigeon.com/cart-abandonment-stats/>
- [17] ConvertCart. (2024, November 19). *Product Page Statistics Every eCommerce Pro Should Know (2025 Update)*. Convertcart. Retrieved January 31, 2025,

- from
<https://www.convertcart.com/blog/ecommerce-product-page-statistics>
- [18] ConvertCart. (2024, November 19). *Product Page Statistics Every eCommerce Pro Should Know (2025 Update)*. Convertcart. Retrieved January 31, 2025, from
<https://www.convertcart.com/blog/ecommerce-product-page-statistics>
- [19] DCatalog. (n.d.). *What is an Ecommerce Product Catalog?* What is an Ecommerce Product Catalog? <https://dcatalog.com/what-is-ecommerce-product-catalog/>
- [20] Experro. (n.d.). *Experro. Product Clustering in Ecommerce*. <https://www.experro.com/blog/product-clustering-in-ecommerce/>
- [21] GitHub. (n.d.). *sohansai/product-categorization: E-commerce Product Categorization Model Using Deep Learning*. GitHub. Retrieved February 1, 2025, from
<https://github.com/sohansai/product-categorization>
- [22] Guide, S. (2024, January 24). *Improve Your Store With Ecommerce Catalog Management (2024)*. Shopify. Retrieved January 31, 2025, from
<https://www.shopify.com/blog/ecommerce-catalog-management>
- [23] Guide, S. (n.d.). *Ecommerce Catalog Management: Tips to Improve + Feedonomics*. BigCommerce. Retrieved January 31, 2025, from
<https://www.bigcommerce.com/articles/ecommerce/catalog-management/>
- [24] guide, s. (n.d.). *Ecommerce Catalog Management: Tips to Improve + Feedonomics*. BigCommerce. Retrieved February 1, 2025, from
<https://www.bigcommerce.com/articles/ecommerce/catalog-management/>
- [25] guide, s. (n.d.). *Ecommerce Catalog Management: Tips to Improve + Feedonomics*. BigCommerce. Retrieved February 1, 2025, from
<https://www.bigcommerce.com/articles/ecommerce/catalog-management/>
- [26] Lopienski, K. (2023, October 26). *Catalog Management: Definition & How to Improve It*. ShipBob. Retrieved January 31, 2025, from
<https://www.shipbob.com/blog/catalog-management/>
- [27] Miva. (n.d.). *Digital Catalog Ecommerce*. Miva. <https://blog.miva.com/digital-catalog-ecommerce>
- [28] nChannel. (n.d.). *Challenges of Ecommerce Catalog Management*. nChannel. <https://www.nchannel.com/blog/challenges-ecommerce-catalog-management/>
- [29] Netscribes. (2023, November 10). *How Catalog AI is Elevating Product Listings in E-commerce*. Netscribes. Retrieved February 1, 2025, from
<https://www.netscribes.com/how-catalog-ai-is-transforming-product-listings/>
- [30] Pimcore. (n.d.). *Product Catalog Management: Why It's Essential for Retail and Commerce*. Pimcore. <https://pimcore.com/en/resources/insights/product-catalog-management-why-essential-for-retail-and-commerce>
- [31] Rishabh Software. (2021, August 13). *Product Catalog Management Software Development Using ML*. Rishabh Software. Retrieved February 1, 2025, from
<https://www.rishabhsoft.com/blog/ml-based-product-catalog-management-software-development>
- [32] Sandelowski, M., & Barroso, J. (2007). *Handbook for Synthesizing Qualitative Research*. Springer Publishing Company, New York
- [33] ShipBob. (n.d.). *What is Catalog Management in Ecommerce? What is Catalog Management in Ecommerce?* <https://www.shipbob.com/blog/catalog-management/>
- [34] Sinha, A., & Gujral, E. (2024, 05 28). *PAE: LLM-based Product Attribute Extraction for E-Commerce Fashion Trends*. arXiv. <https://arxiv.org/abs/2405.17533>
- [35] Sitecore. (n.d.). *What is Ecommerce Catalog Management?* Sitecore. <https://www.sitecore.com/resources/insights/ecommerce/what-is-ecommerce-catalog-management>
- [36] Stanford. (2018). *Attribute extraction from eCommerce product descriptions*. CS229:

- Machine Learning. Retrieved February 1, 2025, from <https://cs229.stanford.edu/proj2018/report/190.pdf>
- [37] TransForm Solutions. (n.d.). *Challenges of eCommerce Product Catalog Management*. TransForm Solutions. Retrieved February 1, 2025, from <https://www.transformsolution.com/blog-posts/challenges-of-ecommerce-product-catalog-management>
- [38] Trustana. (n.d.). *Trustana*. Why Accurate Product Content Matters in E-commerce. <https://www.trustana.com/blog/why-accurate-product-content-matters-in-e-commerce>
- [39] ViSenze. (2023, 10 18). *Cracking the Code of Shopping Cart Abandonment in E-commerce*. ViSenze blog. <https://www.visenze.com/blog/2023/10/18/cracking-the-code-of-shopping-cart-abandonment-in-e-commerce/>
- [40] Vue.ai. (2023, February 28). *Ecommerce Catalog Management : Top 5 Challenges In 2024*. Vue.ai. Retrieved January 31, 2025, from <https://vue.ai/blog/ai-in-retail/ecommerce-catalog-management-challenges/>