

# Sporting Insights Scales Product Categorization Using Generative AI on AWS

Sporting Insights, a premier data and market research firm, manages extensive datasets of product information to provide critical business intelligence. The company partnered with Lambert Labs to build a production-ready, AI-powered categorization pipeline on AWS.

**600k**

products categorized

**2m**

Individual EANs processed

**95%+**

less time spent on manual tasks

The solution leverages Amazon Bedrock to perform sophisticated product categorization using Generative AI, integrated with a scalable serverless architecture featuring AWS Batch and AWS Fargate. This automated workflow enables Sporting Insights to process millions of items efficiently while maintaining high accuracy. By transitioning to this cloud-native pipeline, Sporting Insights has achieved a robust foundation for high-volume data enrichment, significantly reducing manual effort and providing the scalability needed to handle their growing data demands.

Turn over for a deep dive

“ They understood our domain, moved quickly, and delivered a system that genuinely scales with our business. ”

Robert Cobain, Head of Market Intelligence, Sporting Insights

“ Lambert Labs helped us turn what was a manual, human-led process into a production-grade AI pipeline that categorizes tens of thousands of products in hours rather than weeks. ”

Robert Cobain, Head of Market Intelligence, Sporting Insights

## Who are Sporting Insights?

Sporting Insights is a premier research and consultancy firm dedicated to the sports, leisure, and outdoor sectors. They provide global brands and retailers with essential market intelligence and consumer insights to drive strategic decision-making. Through a combination of market research and advanced data analytics, Sporting Insights helps clients navigate complex market landscapes and understand shifting consumer behaviors.



## Spotlight: Amazon Bedrock



Amazon Bedrock simplifies GenAI adoption with secure access to top foundation models via a single API. Rapidly build and scale intelligent applications while keeping your data private and secure.

## Trusted by leading names on AWS



## We are experts in



### AWS Migrations

From on-premise and other cloud providers to AWS



### AWS Modernization

Future-proof your stack with containerization, IaC and serverless



### Data and Analytics

Unlock and turbocharge your data with pipelines and analytics



### AI on AWS

Stay ahead of the curve with GenAI and ML



### AWS Security

Secure your cloud with tailored security reviews



### AWS Managed Services

Place your cloud in the hands of experts

To find out more please contact [sales@lambertlabs.com](mailto:sales@lambertlabs.com)

# Sporting Insights Scales Product Categorization Using Generative AI on AWS

## Opportunity / Customer Challenge

Sporting Insights provides specialized business intelligence by analyzing product data sourced directly from global brands and retailers. A core component of their competitive advantage is the ability to accurately categorize these items into a precise, hierarchical internal taxonomy spanning broad departments down to specific product sub-categories.

As the volume of product data grew from small lists to datasets containing millions of items, the company's reliance on manual categorization became a critical bottleneck. This labor-intensive approach made it impossible to maintain the speed and precision required to exploit their data-driven insights at scale. To sustain their competitive advantage, Sporting Insights needed to move beyond initial ideas and proof of concepts to a production solution. The challenge was to develop a high performance pipeline capable of performing sophisticated data enrichment with human-like domain knowledge while ensuring high accuracy across diverse product lists. The opportunity lay in leveraging Generative AI on AWS to automate the categorization process, allowing Sporting Insights to process massive datasets rapidly and cost effectively while maintaining their same high quality standards.

## Solution

To address the challenge of processing millions of products, Lambert Labs architected a high-scale, event-driven categorization pipeline utilizing Generative AI and serverless compute. The solution is designed to handle massive CSV datasets by automatically partitioning them into smaller, manageable chunks for parallel processing, which ensures high throughput and system reliability. This workflow begins when a product data file is uploaded to an Amazon S3 bucket, triggering an AWS Lambda function that handles the data chunking. These segments are then managed via Amazon Simple Queue Service, which acts as a buffer to regulate the flow of work and prevent system overload.

For the complex task of categorization, the system utilizes Batch and Fargate to execute containerized jobs that interact with Bedrock. By leveraging foundational Large Language Models (LLMs) via Bedrock, e.g. Anthropic's Claude Sonnet 4, the pipeline applies sophisticated reasoning to assign products to the correct internal taxonomy based on customized prompts and detailed examples. This model-driven approach allows the system to replicate human domain knowledge at a scale previously impossible with manual methods.

To ensure the solution remained cost effective and efficient, the architecture incorporates intelligent caching and rigorous model selection. A comprehensive model comparison was conducted to justify the choice of LLM based on both accuracy and price, while Amazon DynamoDB was implemented to cache external API requests and reduce redundant processing costs. Once the AI has categorized the data, the final output files are written back to an S3 bucket, where they can be easily accessed and utilized by Sporting Insights' analytics team for their ongoing business intelligence needs.

## Outcome

The implementation of the AI-powered categorization pipeline has successfully transformed how Sporting Insights manages and enriches its massive datasets. By migrating from a manual, labor-intensive process to fully automated AWS infrastructure, the company has eliminated a significant bottleneck. The new system allows Sporting Insights to process millions of products with a level of speed and consistency that was previously unattainable, ensuring that their specialized business intelligence remains accurate even as data volumes continue to surge.

Ultimately, this project has solidified Sporting Insights' competitive edge in the market. With a production pipeline now in place, the company can rapidly ingest data from global brands and retailers, categorize it according to their precise internal taxonomy, and deliver actionable insights with unprecedented efficiency. This robust foundation not only meets their current processing demands but also provides the long-term scalability required to support their evolving business objectives and expanding datasets.

Trusted by leading names on AWS



BALLOTPEDIA

To find out more please contact [sales@lambertlabs.com](mailto:sales@lambertlabs.com)