

Case study

ProductMatch MVP

Data-Driven Methodologies for Pinpointing High-Yield Products on Amazon

Technologies & frameworks



Django



Scrapy



TensorFlow



Python



C++



C



Celery



Redis



ProductMatch

Client since
2023

Location
United Kingdom

Industry
E-commerce

Clients stats

Shops
750K

Products
1.7B

Crawling power
100k products / h

The screenshot displays the ProductMatch MVP interface. At the top, there's a navigation bar with a 'Shopify' dropdown and three buttons: 'Main' (hex #276FBF), 'BG' (hex #EBF3F8), and 'Shade' (hex #BBDEFB). Below this is a header for 'PlanetX' with fields for 'URL ADDRESS' (planetx.co.uk), 'WEBSITE NAME' (PlanteX - #13), 'TYPE' (Shopify), and 'LAST CRAWLED' (31 Oct 23). A search bar and buttons for 'Start crawling' and 'Start matching products' are present. The main content is a table with columns: Amazon ID, Current Price, Current Rank, Current Cat., FBA Fee, All Time Avg. Rank, Avg. 90 Days Price, Monthly sold avg. all time, ROI %, Profit Margin %, and Similarity. A 'Product from your website' section shows a 'Current match' for 'KMS AddVolume Shampoo 300ml Double' with a price of £34.80. A modal window titled 'Create a new website for crawl' is open, showing a form with 'PlanteX' as the website name, 'Shopify' as the type, and a 'URL Address..' field. The modal has 'Cancel creation' and 'Create website' buttons. A 'Home' button is visible at the bottom right.

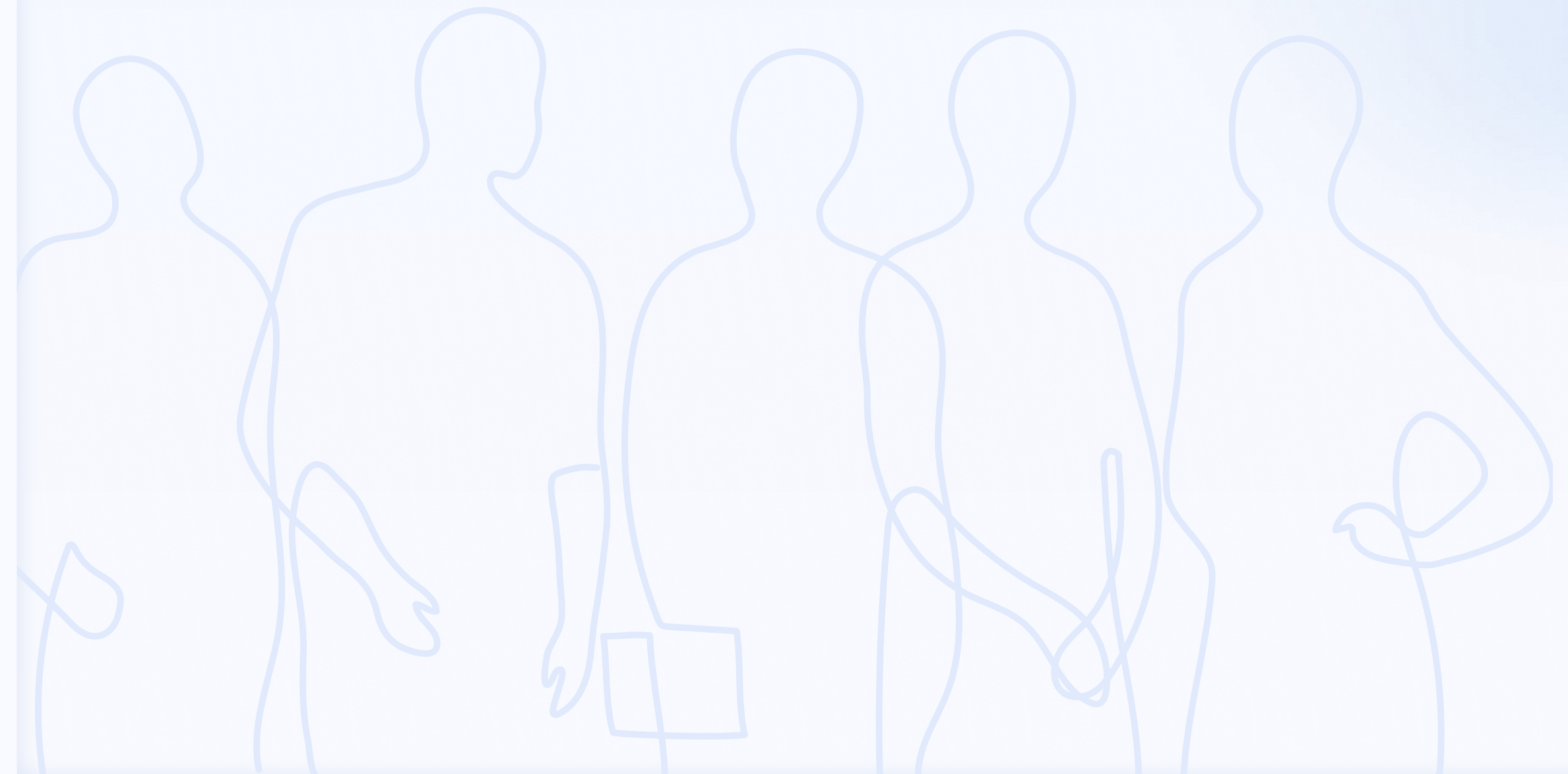
About the client

Our client is a distinguished third-party vendor on the Amazon platform, specializing in strategic market analysis.

Through rigorous digital marketplace research, they identify high-performing products and significant discount opportunities.

Their approach is twofold:

- Introducing unrepresented products to the Amazon landscape.
- Competitively pricing offerings to ensure a distinct market advantage over existing listings.



The Challenge

⚠ In the competitive world of Amazon third-party selling, our client faced substantial challenges. Their strategy, rooted in manual searches across countless websites for lucrative Amazon products, highlighted the pressing need for an integrated digital solution.

Relying heavily on various SaaS tools to retrieve Amazon product histories, they sought a streamlined system to consolidate and analyze this data efficiently.

Their operational challenges were exacerbated by the lack of a system that could seamlessly pair products with their Amazon equivalents, quickly assess ROI, and guarantee up-to-the-minute price accuracy.

Implementation

- + Development of Generic Crawling Algorithms specifically for Shopify, WooCommerce and SEO optimised sites.
- + Crafting of Amazon-Specific Crawling.
- + Deployment of AI-Powered Product Matching using pre-trained e-commerce models.
- + Introduction of an Intuitive UI with admin interfaces.
- + Utilization of Django Orchestration complemented by Celery tasks.
- + Seamless Scrapy Integration within Django.
- + Incorporation of a TensorFlow-Powered Microservice for product matching.
- + Establishment of a Redis & Celery Partnership.
- + Implementation of the WebIdentity Mechanism.
- + Introduction of a Captcha Cracking engine.

The Results

- ✓ Access to data from over 750,000 UK websites through tailored crawling.
- ✓ Constant updates on Amazon product pricing data.
- ✓ Precise matches between scraped products and their Amazon equivalents.
- ✓ Clear view of products with high profitability potential on Amazon and comprehensive oversight of all processes.
- ✓ Efficient asynchronous operations using Django and Celery.
- ✓ Proficient crawling of Shopify, WooCommerce and SEO optimised sites.
- ✓ Efficient inter-service communication for product matching via REST.
- ✓ Adept task management through Redis & Celery.
- ✓ Reduced need for an extensive proxy pool due to genuine request simulation.
- ✓ Optimized Amazon crawling rates through captcha cracking.