

# FashionBrain Project: A Vision for Understanding Europe's Fashion Data Universe



Alessandro Checco, Gianluca Demartini, Alexander Löser, Ines Arous, Mourad Khayati, Matthias Dantone, Richard Koopmanschap, Svetlin Stalinov, Martin Kersten, Ying Zhang

University of Sheffield, Sheffield, UK; Beuth Hochschule für Technik Berlin, Germany; University of Fribourg, Switzerland; Fashwell AG, Switzerland; MonetDB Solutions, The Netherlands



## Introduction

In the world of fashion, retailers often either do not own enough data to predict customers' next trends or these data is not integrated in a way that can create valuable insights.

In order to alleviate the existing dependence from social networks and search engines, fashion retailers should be able to use their own tools and data to predict emerging trends, and to acquire fashion related data by other means, for example by crowdsourced activities or by tailored user interactions.

## Acquisition Channel

Figure 1 illustrates the current and envisioned acquisition channels for an online-fashion shop. Most traffic of the online shop is currently forwarded from major search engines. The shop owner 'buys' this traffic from search engine owners with advertisements displayed on the search engine, or with investments in search engine optimizations of their own portal. Much fewer potential customers enter the shop directly navigating to the shop website and very few customers enter the shop via links from third party web pages, such as a blog from a fashion influencer.

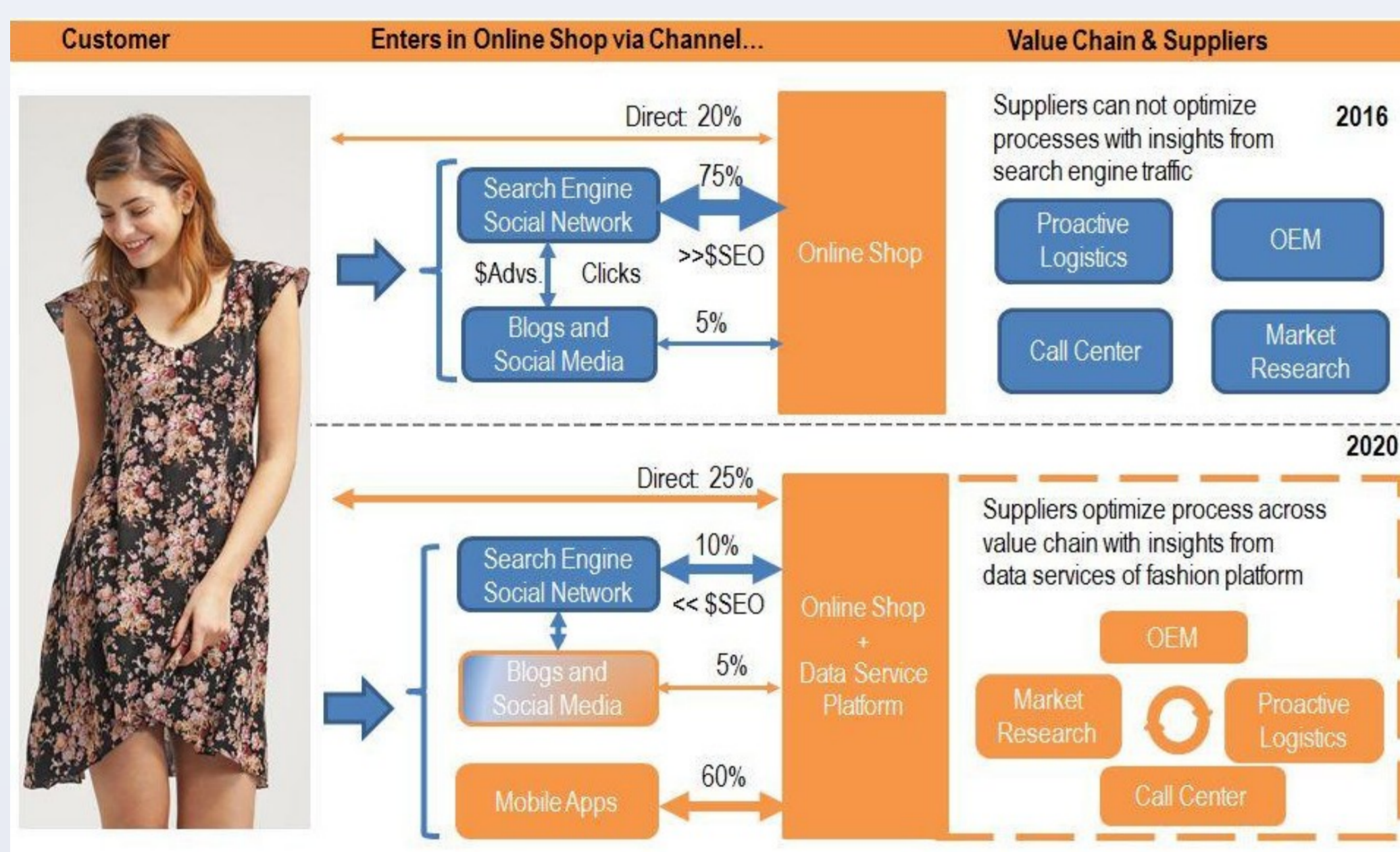


Figure 1: Typical and envisioned acquisition channels in online shops.

## Use Case – Shop the Look

On top of our integrated fashion data platform, we envision end-user applications like the one presented in Figure 2, where the user is issuing a query consisting of an image (e.g., taken with a mobile device). The proposed system takes the input image and, thanks to a mix of image processing algorithms and crowdsourcing workflows, performs extraction of fashion products from the image followed by entity linking against a retailer product catalog that allows to disambiguate similar products. This allows the user to identify the exact product depicted in the image and enables him to purchase the product directly. More than search, the system recommends additional products (e.g., an hat) that is not depicted in the image but fits well with the identified products.

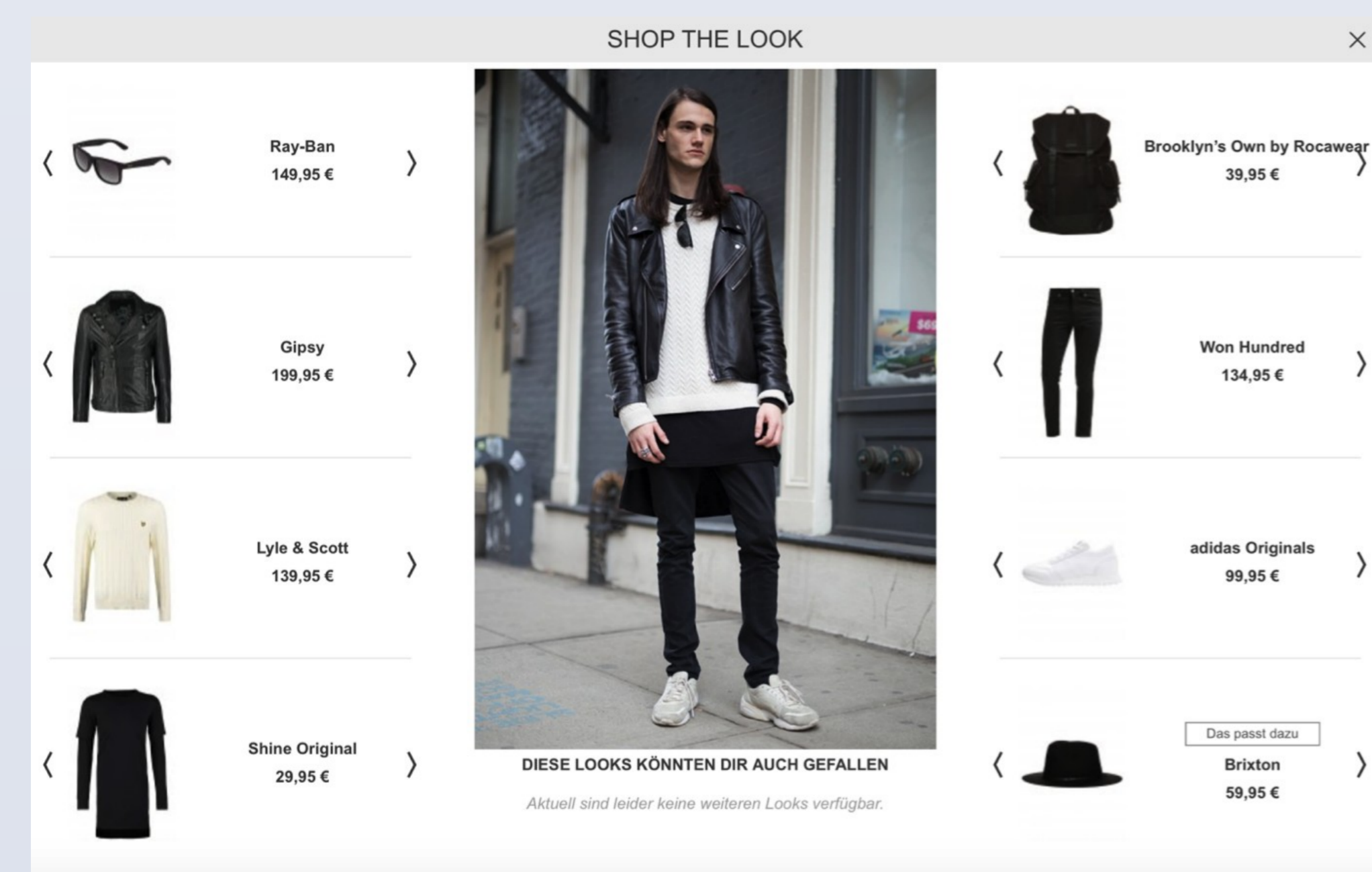


Figure 2: Example of photo-based product search with catalog matching and recommendation.

## Use Case - Fashion Trend Prediction

Using similar hybrid human-machine approaches used for the previous use case where we combine image processing algorithms with human computation by means of crowdsourcing, we envision a system that takes as input Instagram images from a list of well know fashion bloggers. These data will be then processed over our pipeline that links depicted fashion products to retailer product catalogs. Then, thanks to time series analysis techniques that embed social media with sales data, the system will be able to *predict upcoming fashion trends*. Such predictions will enable retailers to better manage the supply chain and make sure that product stock is always up to date with expected demand.

## Challenges

As summarized in Figure 3, we seek an integrated approach for textual, structured, multimedia and longitudinal data.

We focus on the following three core topics:

- **Data processing.** The collection of heterogeneous data, its integration and curation is a problem broadly studied in the literature. However, in the world of online fashion retail, typical high performance ad-hoc solutions are required. The main problems in this area are the integration amongst different data infrastructures and sources (e.g., from retailers, manufacturers, social media, logistics, website, customer care, etc.) and the complexity of the workflow needed to enable complex queries over available integrated data.
- **Predictive analytics.** In the world of fashion the main challenges we identified are the scalability of predictions and the lack of training data to build supervised models.
- **Data management.** The practical implementation of high performance, state-of-the-art data management solutions is a challenge in the world of Big Data: the main direction we explore in our project is the development of low level, in-databases functionalities in the context of memory resident databases.

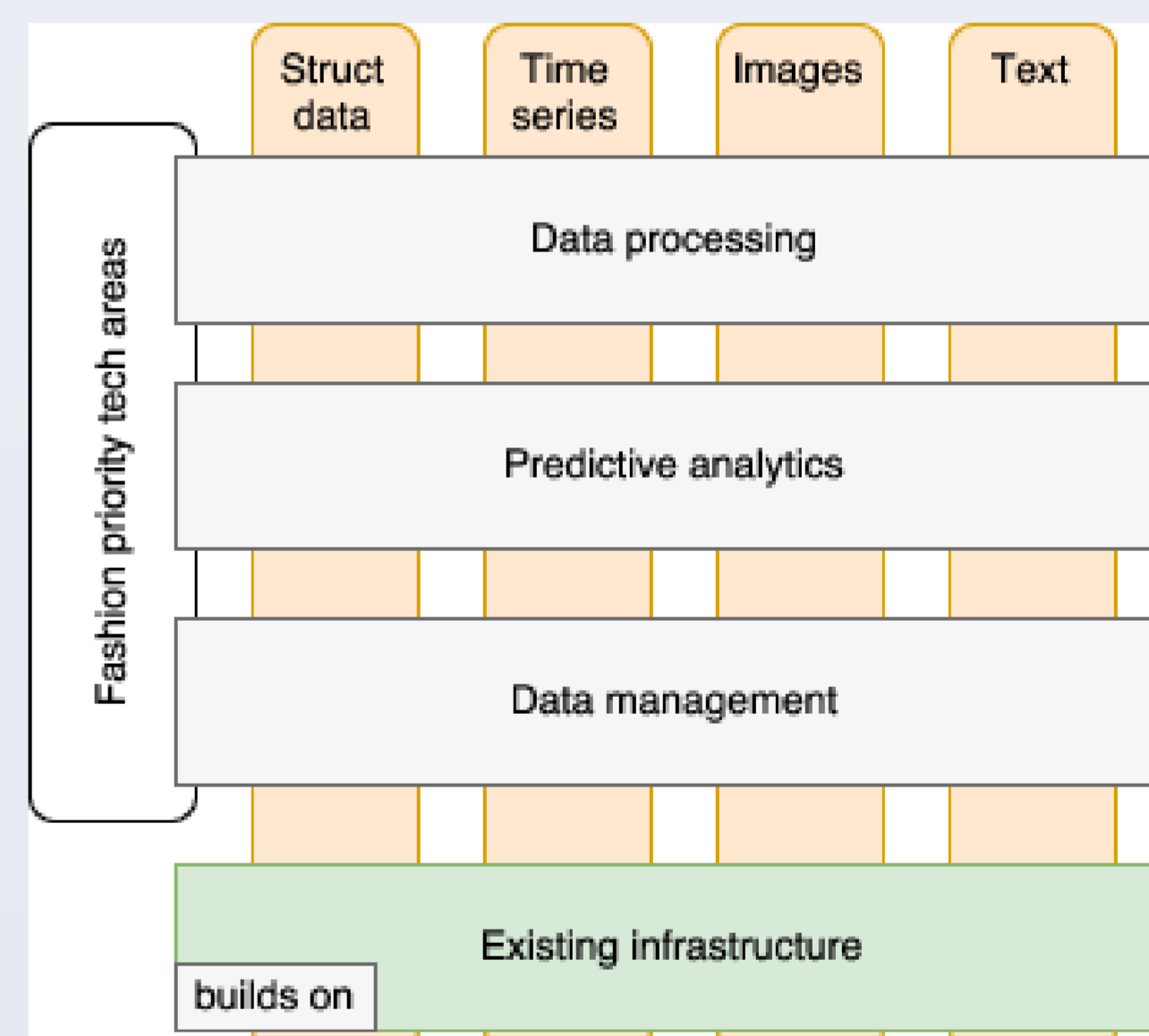


Figure 3: Fashion Big Data reference model.

## Integration Solution

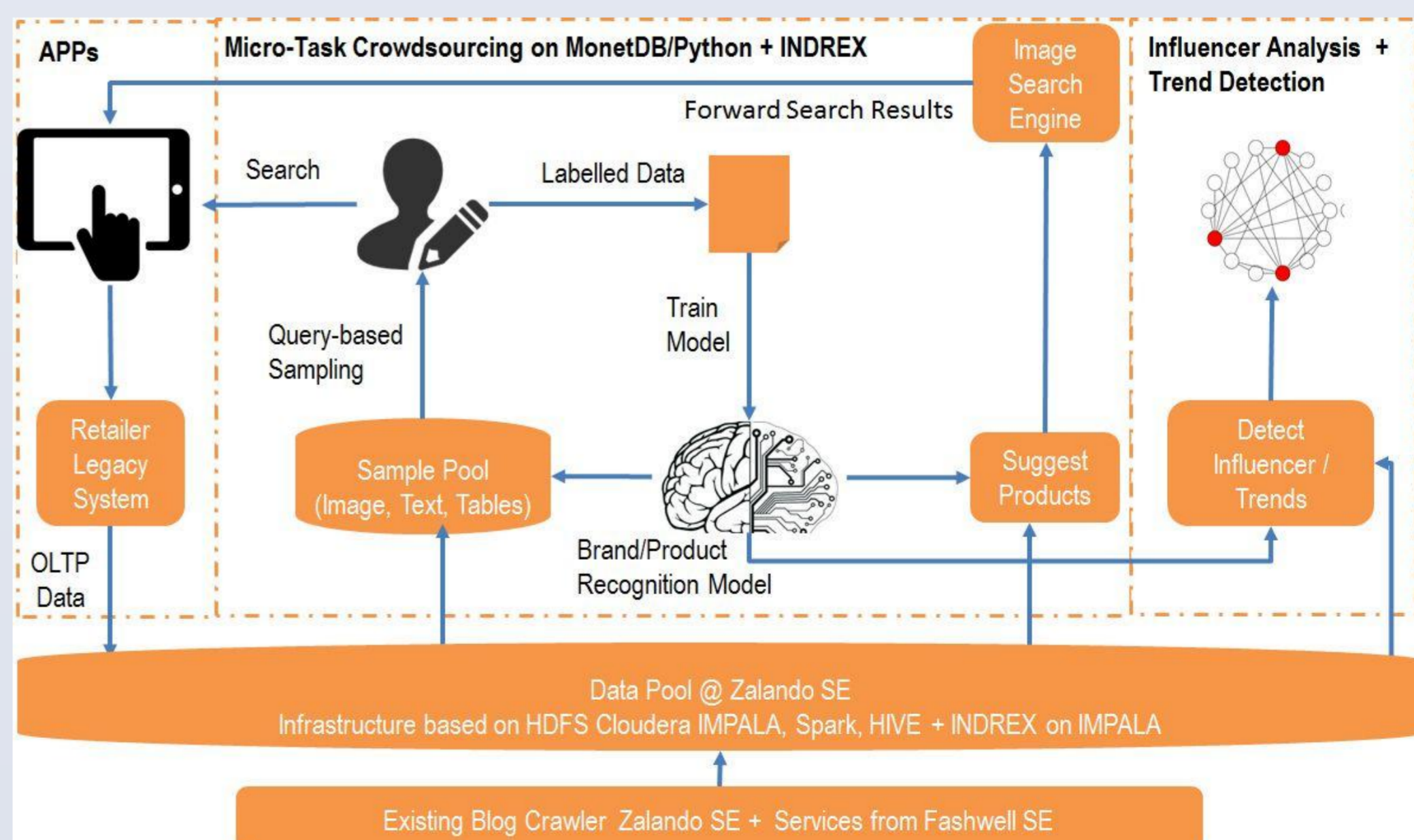


Figure 4: System Architecture.

## Contacts

a.checco@sheffield.ac.uk, g.demartini@sheffield.ac.uk  
This project is supported by the European Union's Horizon 2020 research and innovation programme under grant agreement No. 732328.