



Mittuniversitetet

MID SWEDEN UNIVERSITY

Integrating Semantic with Keyword Product search in Online Marketplaces

By: Roberto Piran Amedi & Albin Rönnkvist



Overview

1. Introduction
2. Literature Review
3. Results and Discussion
4. Proposed Solution
5. Conclusions and Future Recommendations
6. Questions



Introduction

Importance of online marketplaces and product search

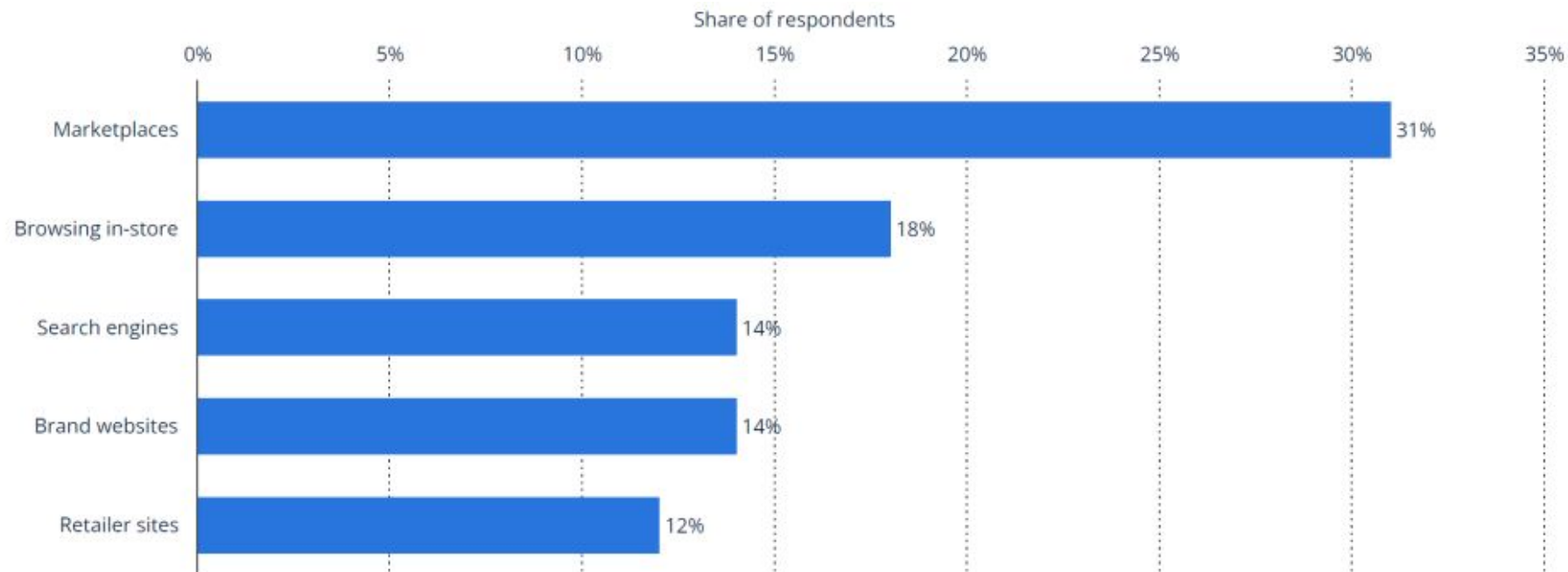


Mittuniversitetet

MID SWEDEN UNIVERSITY

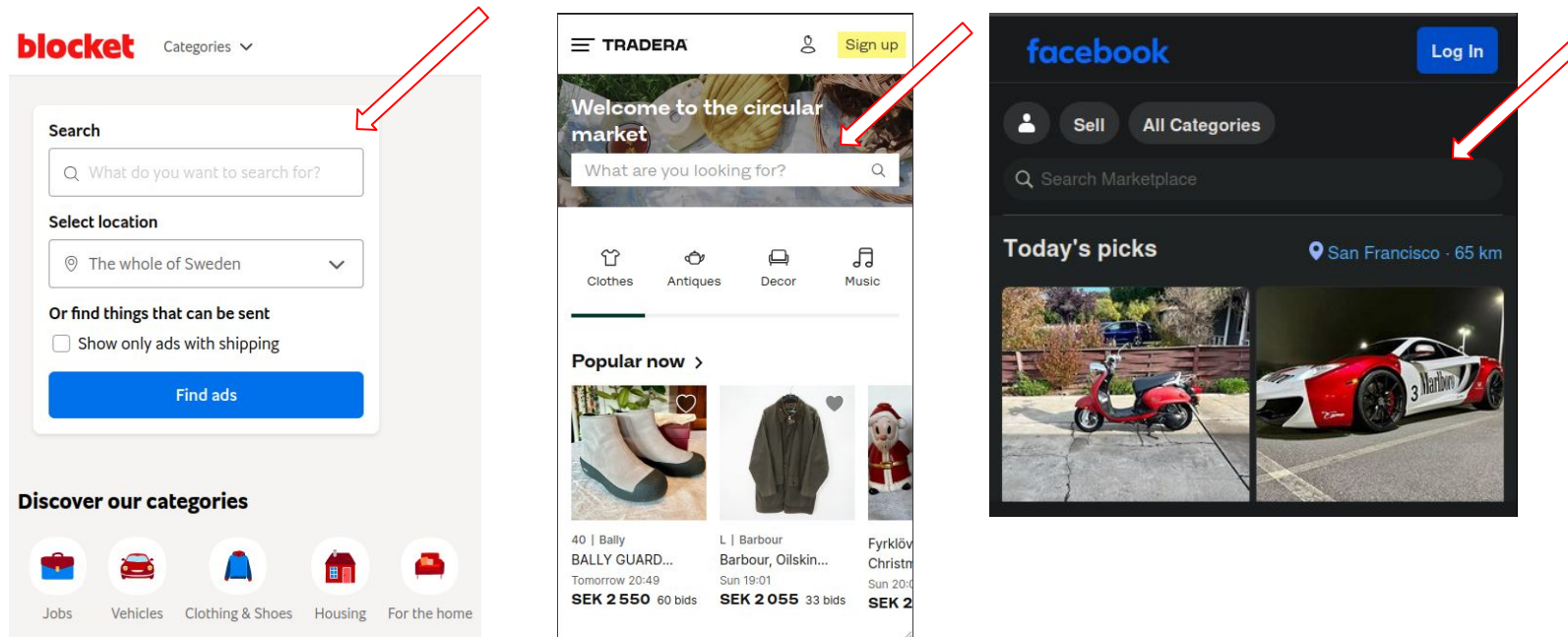
Introduction | Online Marketplaces

- Online marketplaces are popular platforms for finding products
- *2023 International Trade Administration Survey*: Online marketplaces are the main starting point for product searches worldwide



Introduction | Product Search

- The search bar is a widely used tool for finding products
- *2022 Google Survey*: 7 out of 10 shoppers use the search bar as the primary method





Introduction

Current state and challenges of product search

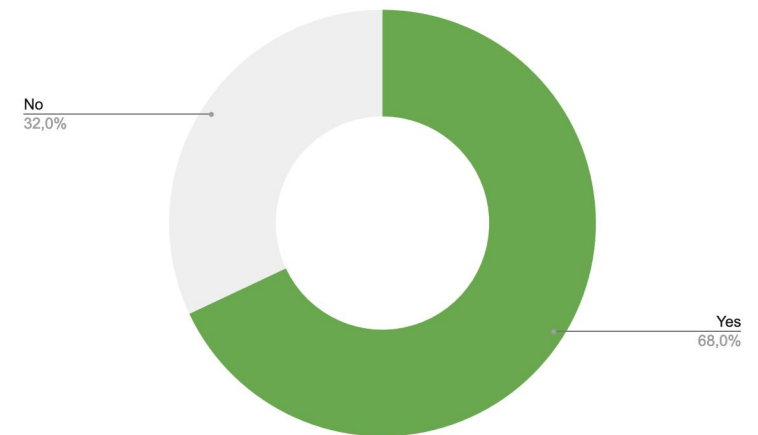
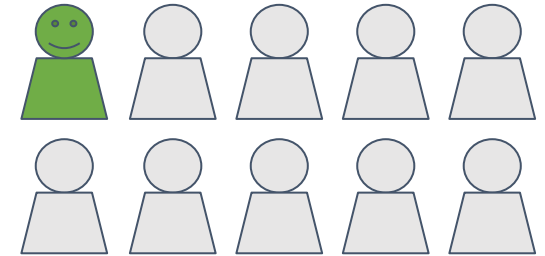


Mittuniversitetet

MID SWEDEN UNIVERSITY

Introduction | Current State

- There is widespread dissatisfaction with search results among shoppers
- *2022 Google Survey*: Only 12% of shoppers find exactly what they are searching for when using the search bar
 - Results in \$2 trillion in annual losses globally
- *2024 Constructor Survey*: 68% of shoppers want improved search results



Introduction | Challenges

- Diverse user intentions

The screenshot shows a search results page for "iphone 16 pro max case". The search bar at the top is highlighted with a red box. Below the search bar, there are filters for "In-store", "Price", "Brand", "Fulfillment Speed", and "Sort by | Best Match". The main content area displays "Results for 'iphone 16 pro max case' (290)". Below this, there are four product cards, each with a heart icon, a "Options" button, and a price. The first card shows a black case for \$13.99. The second card shows a black case for "Now \$17.99" (originally \$29.99). The third card shows a pink case for \$14.99. The fourth card shows a black case for "See price in cart".

The screenshot shows a search results page for "Gift for my 4 year old son". The search bar at the top is highlighted with a red box. Below the search bar, there are filters for "In-store", "Price", "Brand", "Fulfillment Speed", "Availability", "Special Offers", "Customer Rating", "Occasion", and "Gender". The main content area displays "Top types of 'Gift for my 4 year old son'". Below this, there are three product cards, each with a heart icon, a "Best seller" badge, and a "+ Add" button. The first card shows a red toy car for \$14.99. The second card shows a white toy car for "Now \$156" (originally \$8.99). The third card shows a red toy car for "Now \$16.99" (originally \$29.99). The fourth card shows a JoyStone 34 in 1 Dinosaurs Truck, Dino Transport Car... for \$29.99.

Introduction | Challenges

- Diverse user base
 - Language and culture
 - Backgrounds and expertise
- Formulating queries
- Formulating product descriptions

Experienced computer user

16-inch laptop, 16GB RAM, 512GB SSD

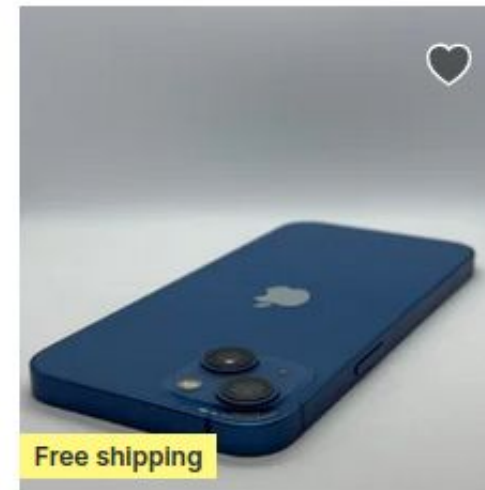


Inexperienced computer user

laptop with big screen and very fast



iPhone



Free shipping

Apple phone device

44 min

SEK 4 599 Or buy now SEK 6 099



Mittuniversitetet

MID SWEDEN UNIVERSITY



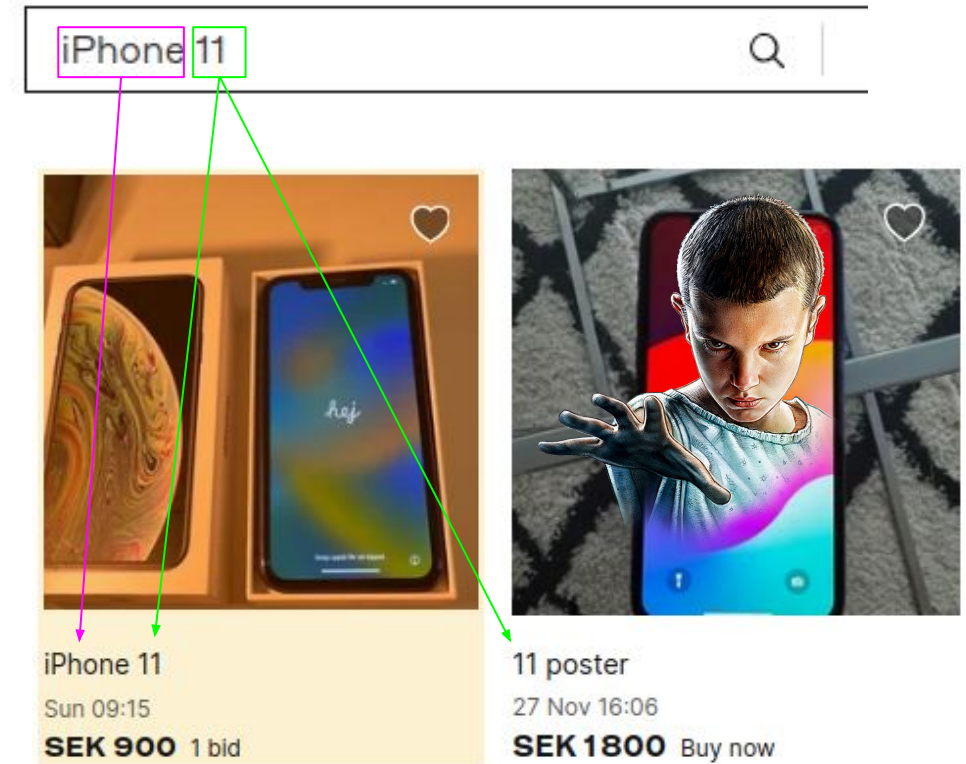
Introduction

Common product search systems: keyword-based, semantic-based, and hybrid



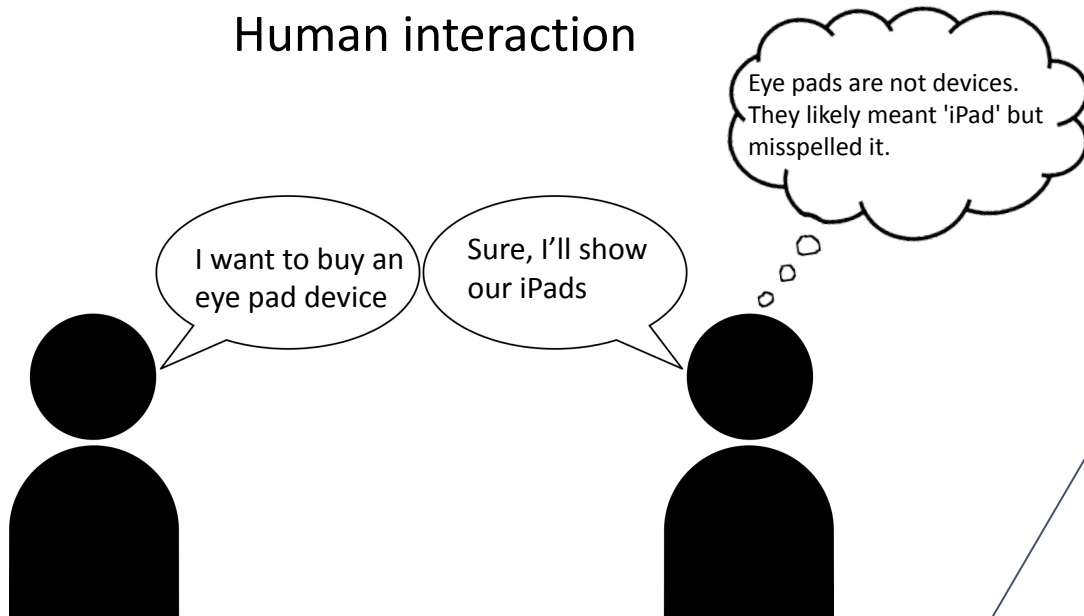
Introduction | Keyword Search

- Word matching
- **Main strength:** Precise matching (e.g., model numbers and brand names)
- **Limitations:**
 - Lacks understanding of context and underlying intent
 - Struggles with vocabulary gaps (e.g., misspellings and synonyms)

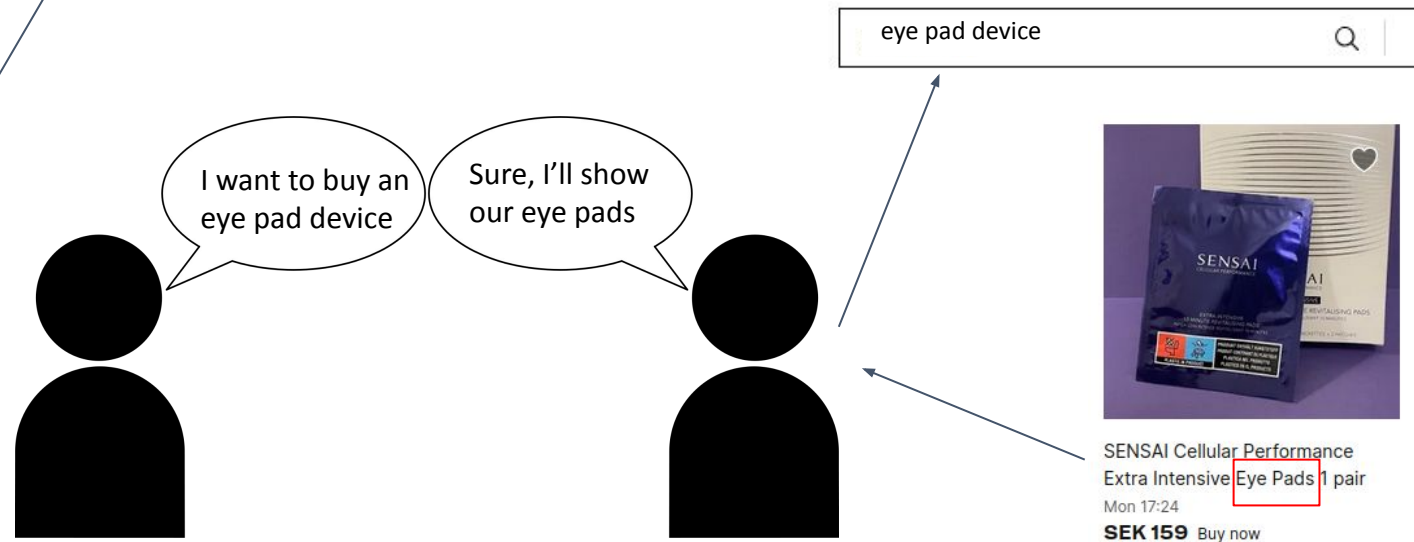


Introduction | Keyword Search

Human interaction

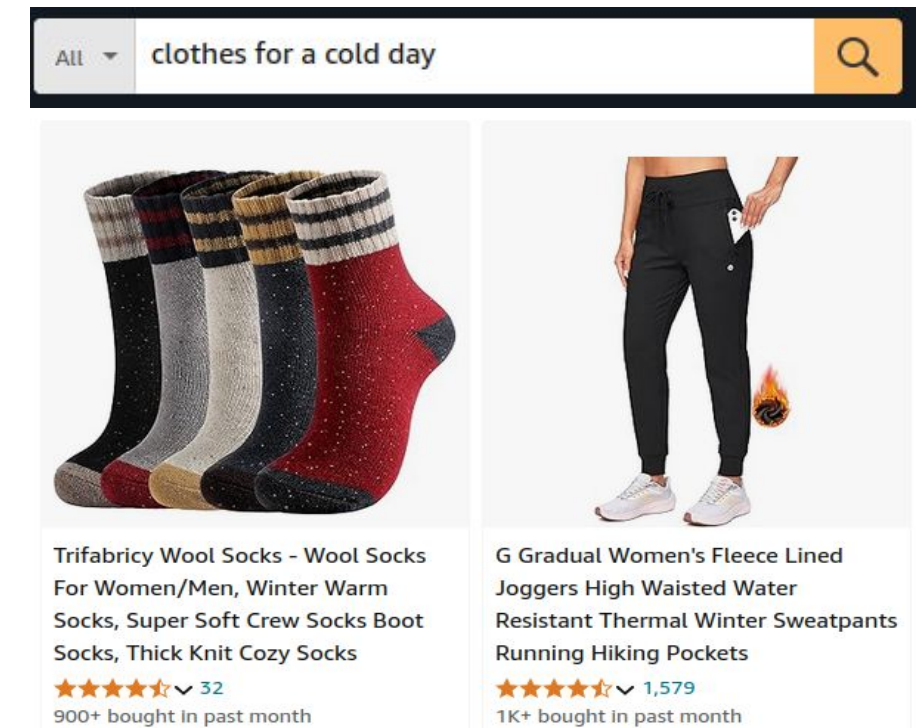
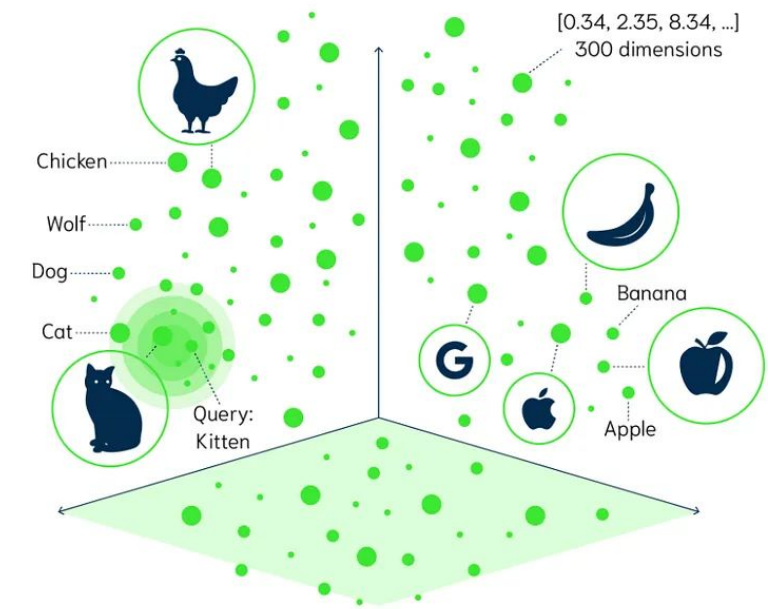


Keyword-based search system



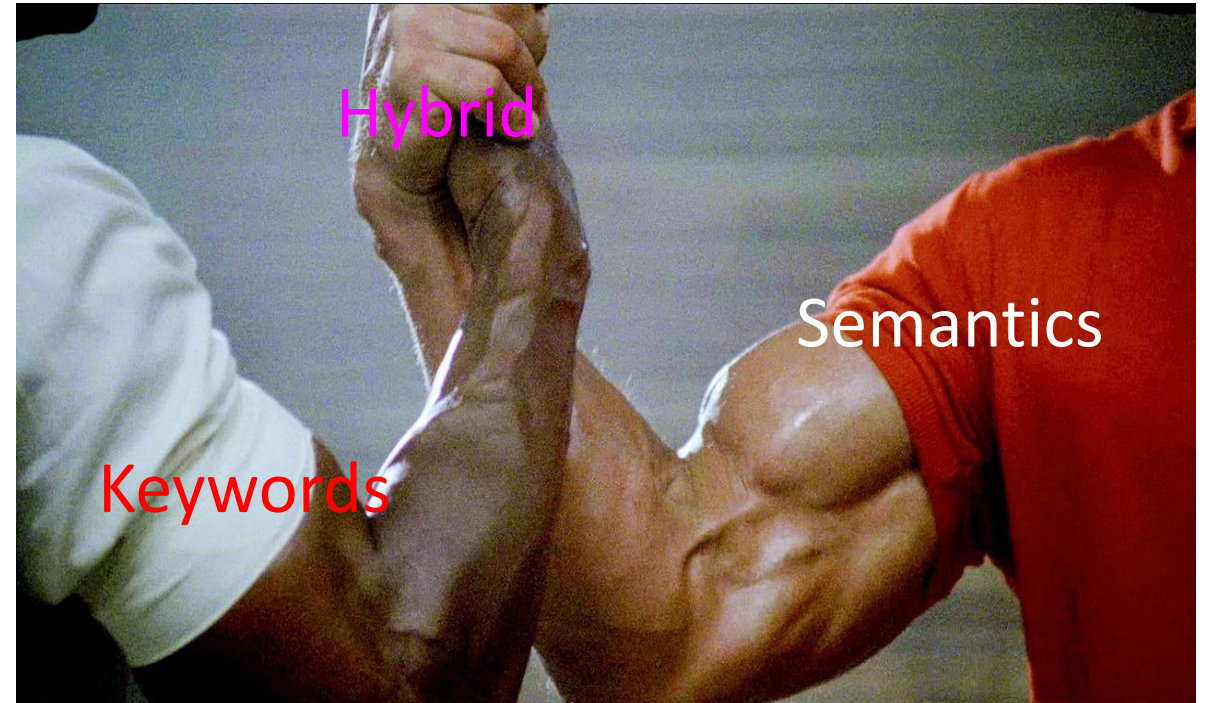
Introduction | Semantic Search

- Matching based on the intent and contextual meaning
- “Product / query” -> Embedding model -> [0, 7, 2, ...]
 - Captures semantic meaning
 - Measure distance or similarity
- **Main strength:** Understands the underlying meaning, beyond just words
- **Limitations:**
 - May produce inaccuracies when specific terms are important (e.g., model numbers)
 - Other known limitations, such as negations and units of measurement



Introduction | Hybrid Search

- Combines keyword and semantic search
- Complementary:
 - Precise term matching
 - Underlying meaning





Introduction

Research Questions



Introduction | Research Questions

- **Hybrid Architectures:** How can hybrid models be implemented to utilize the capabilities of both keyword and semantic models?
- **Hybrid Search Impact on Product Search Results:** What is the impact of hybrid models on product search results compared to keyword and semantic models used independently?





Literature Review

Selected studies, hybrid architectures and their impact on product search results

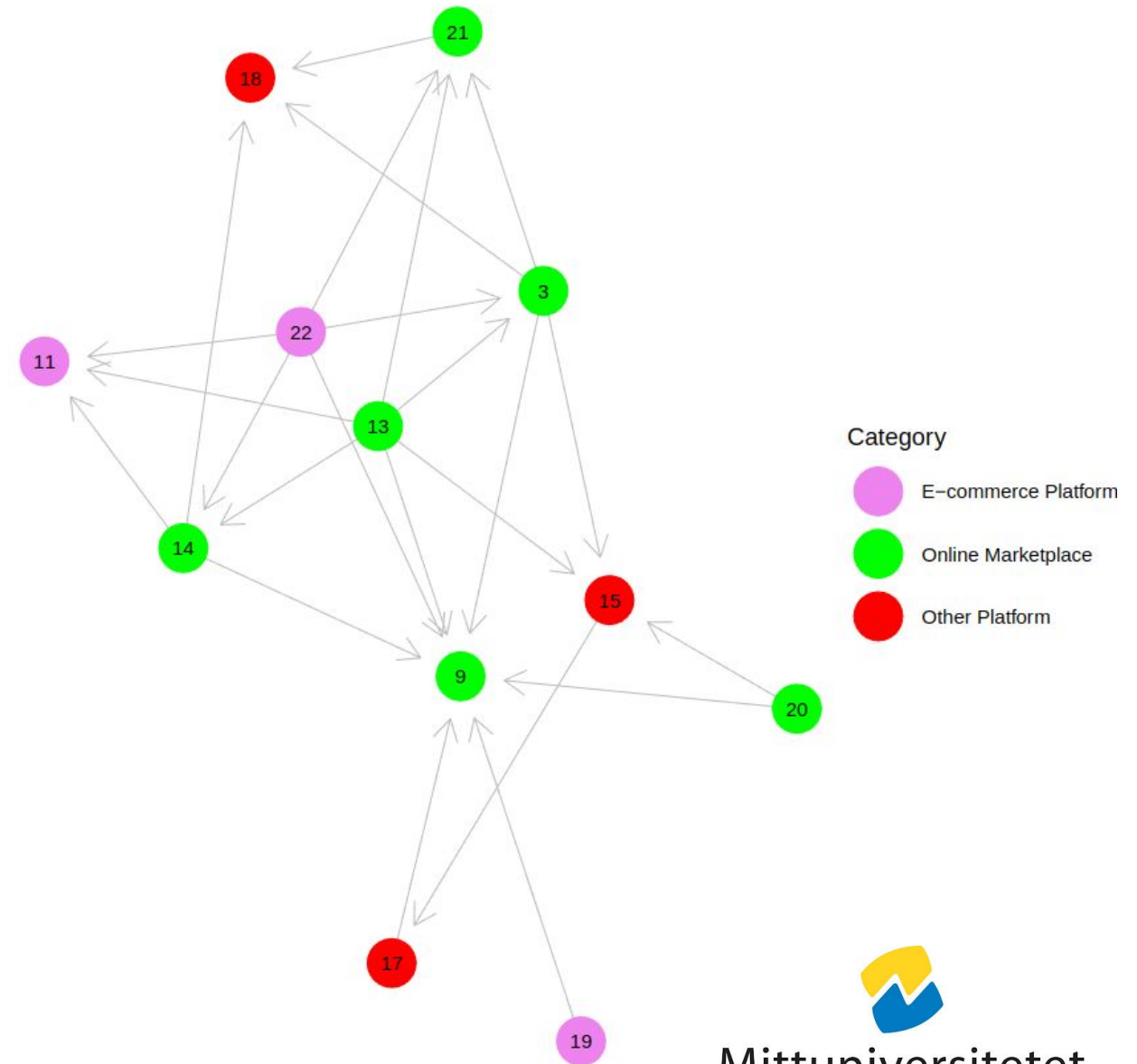


Mittuniversitetet

MID SWEDEN UNIVERSITY

Literature Review

- Selected studies:
 - 6 online marketplaces
 - 3 e-commerce platforms
 - 3 other platforms
- Includes:
 - Hybrid models
 - Offline and online experiments
- The studies are interconnected





Results and Discussion

Hybrid architectures and their impact on product search results

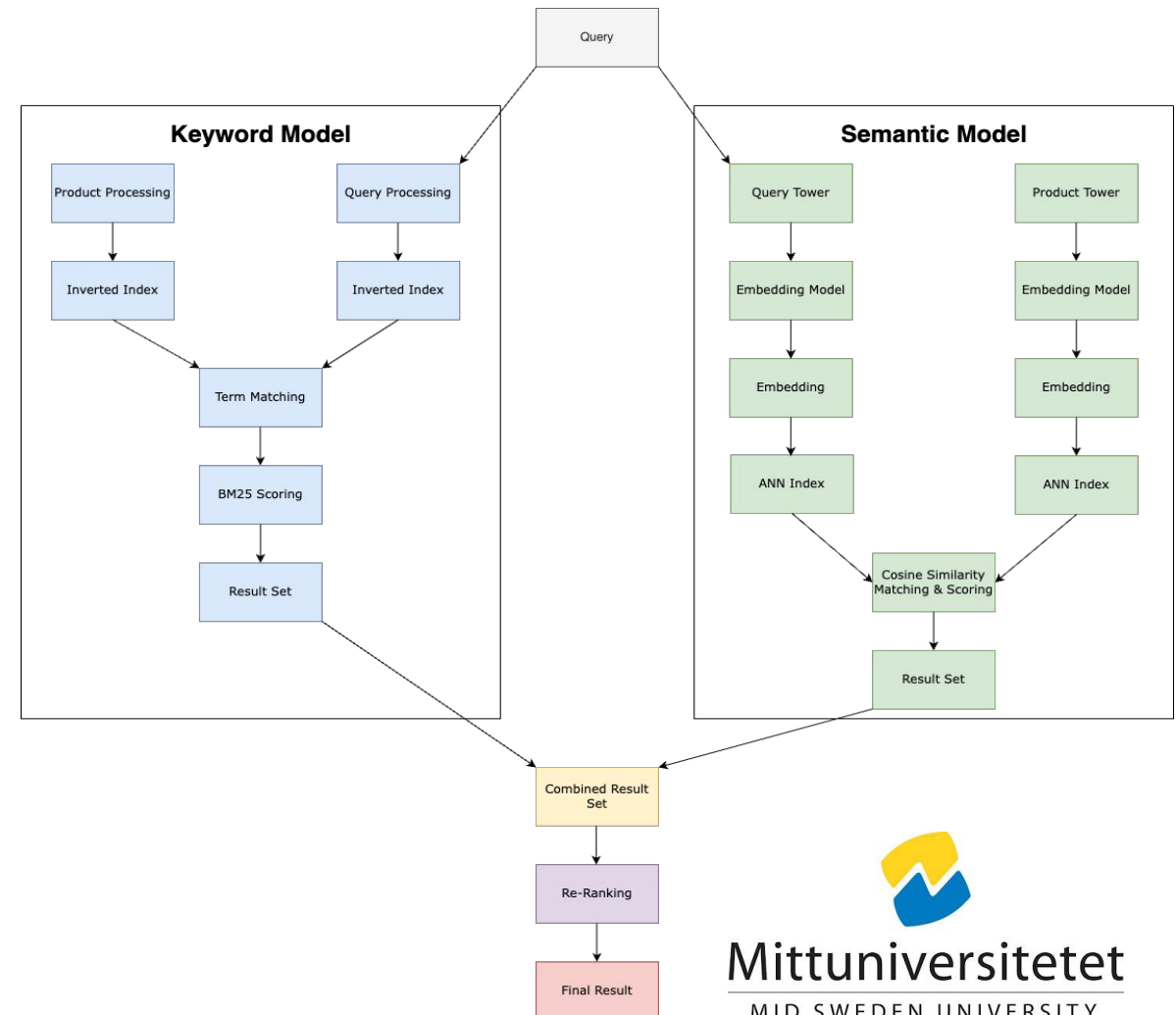


Mittuniversitetet

MID SWEDEN UNIVERSITY

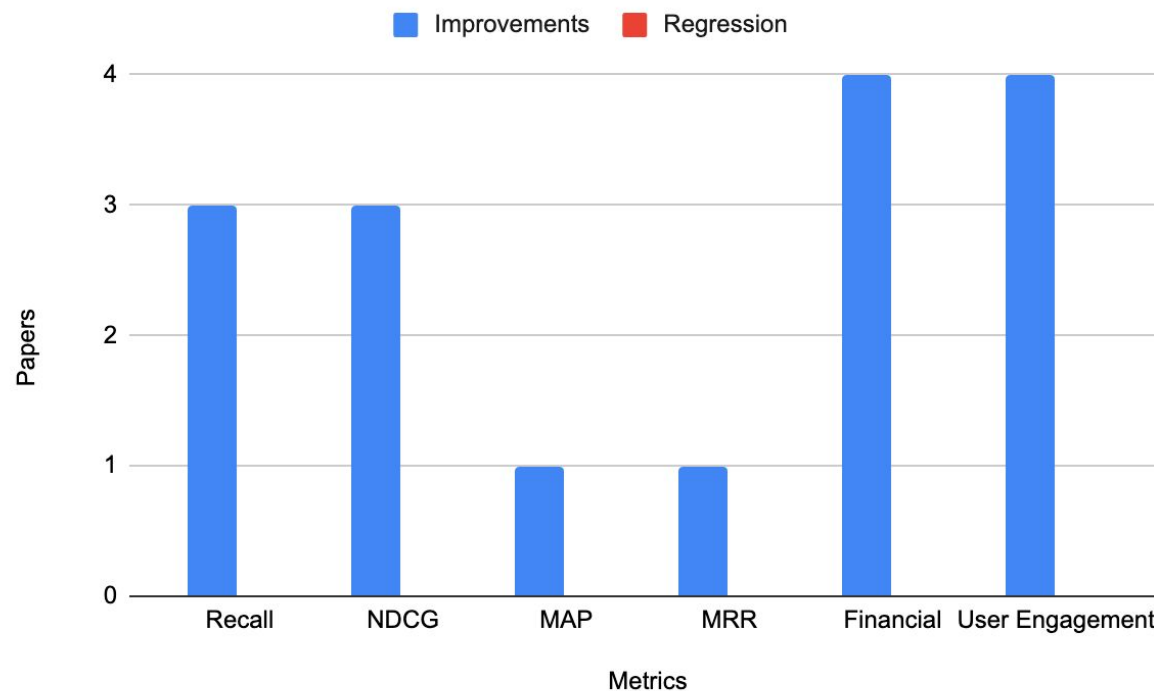
Results and Discussion | Hybrid Architectures

- Differences between hybrid models, but similar foundation
- **Core Keyword Models:**
 - Inverted index term matching
 - BM25 scoring
- **Integration of Semantic Models:**
 - Two-tower architecture with query and product embeddings
 - Approximate Nearest Neighbor (ANN) search with cosine similarity for matching and scoring
- **Hybridization:**
 - **Merging:** Combines results from both models into a unified match set
 - **Re-ranking:** Following this merging, refines and ranks matches based on relevance (limited re-ranking details available across studies)



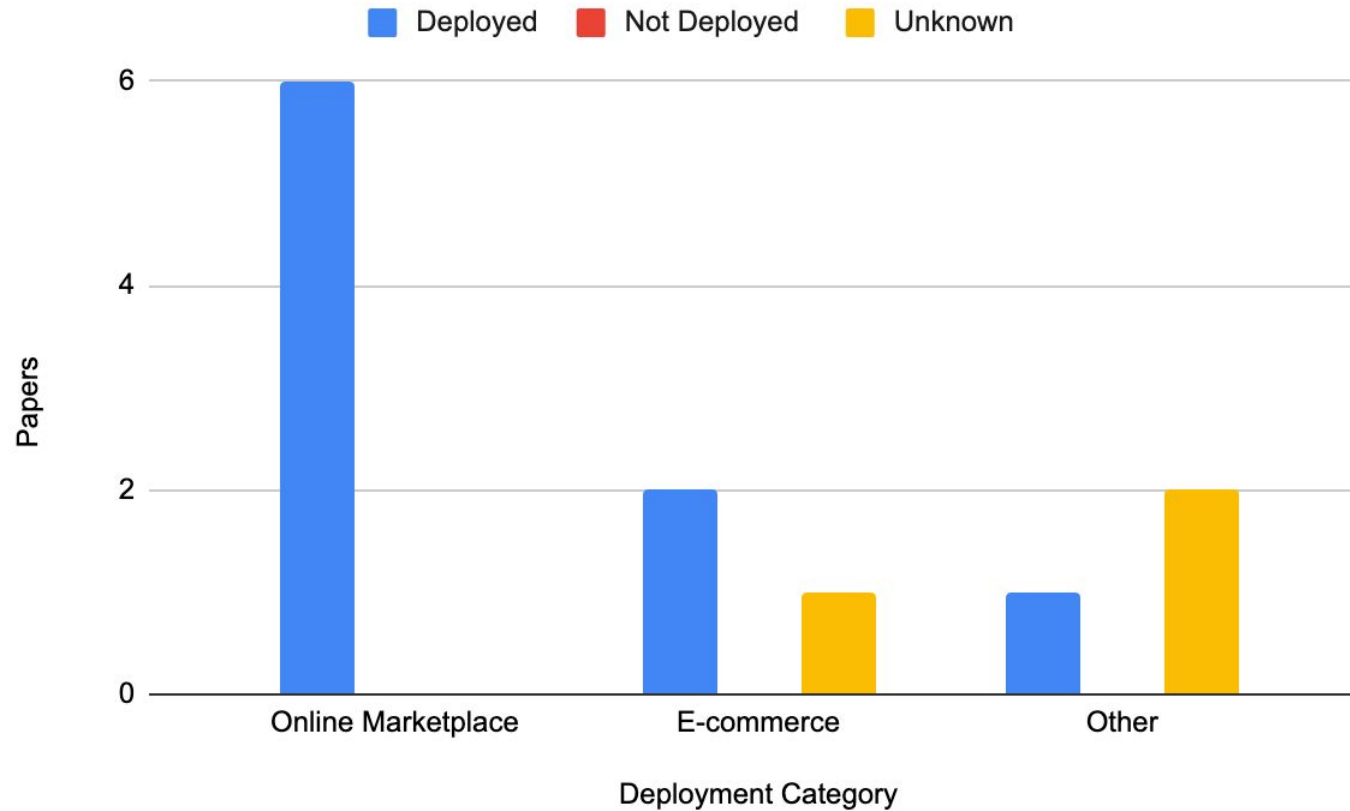
Results and Discussion | Impact

- Significant improvements in product search result metrics when compared to standalone models
- Improved coverage, ranking quality and business performance



Results and Discussion | Impact

- A majority of hybrid models were deployed in production

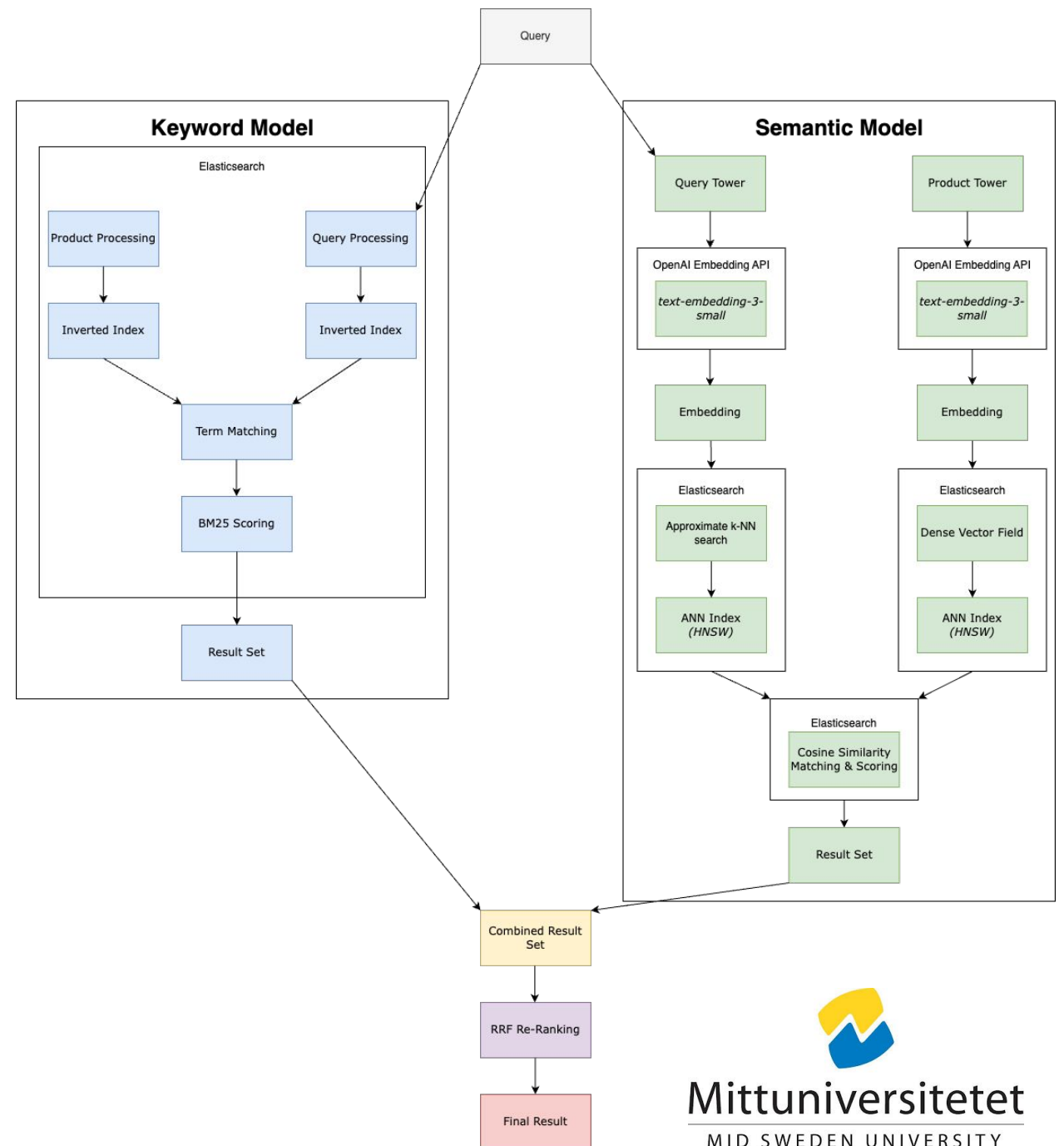




Proposed Solution

Proposed Solution

- Hybrid models provide better product search results
- Hybrid models can be complex to implement
 - Requires expertise and resources
- **Proposed solution:** A simplified hybrid model for broader adoption
 - Inspired by models in the review
 - Lightweight and flexible
 - [Codebase](#)



Conclusions and Future Recommendations



Conclusions and Future Recommendations

- Discussed different hybrid architectures and identified a similar foundation across the literature
- Multiple studies, with various experiments and metrics, verified that hybrid models significantly enhance product search results
- We proposed a simplified hybrid model as a starting point
- Suggest further exploration of the overall hybrid model impact:
 - Complexity and maintenance
 - Resource requirements
 - Cost implications
 - Latency
 - Pagination



Mittuniversitetet

MID SWEDEN UNIVERSITY

Questions