



# In-Session Personalization

## Hands-on Workshop on ML for Retail



Jacopo Tagliabue

Lead A.I. Scientist

Toronto  
Machine Learning  
Society (TMLS)



Machine Learning in Retail Summit

April 21st

# Agenda

- 1. Pleased to meet you!**  
Hope you guessed my [name!](#)
- 2. From personalization to in-session personalization**  
Why sessions are the most important partition for our models
- 3. (Product) space, the final frontier**  
How to build product embeddings, and why they are so important
- 4. One representation to rule them all**  
How session representation can enrich all kinds of models
- 5. Hands-on coding**  
“Talk is cheap, show me the code”
- 6. What’s next?**  
The Future is not what it used to be

# Pleased to meet you!

# 360° relevance in **Commerce**, **Service**, **Website** and **Workplace**



## coveo



# \$325M

Capital raised since 2018  
*for R&D, growth  
and acquisitions*



# 150

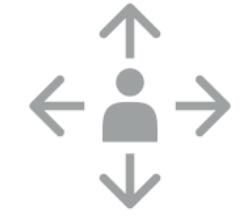
Certified SI Partners  
And Integrations  
*+strategic alliances  
& integrations with key ISVs*



# Global

7 offices, 4 Data Centers  
around the world

*for scale, performance,  
local compliance*



# 1,500

Customer Deployments  
*globally & across  
multiple use cases*



# #1

US/Gartner/Forrester leader  
FORRESTER® **Gartner**

*#1 applied AI platform  
company in Canada*



# ROI

Proven  
Customer Success  
*measured  
business value*

## Refine by

 For Home For Work

## Product Type

 Accessories 11

## Price

 Under \$500 2 \$500 to \$1,000 5 \$1,000 to \$2,500 2 \$2,500 to \$5,000 1 \$5,000 and more 1

## Display Size

 65 inches and more 2 49 to 60 inches 2 32 to 48 inches 1 24 to 32 inches 5 24 inches 1

## Category

 LED-Backlit LCD Monitor 7 LED-Backlit LCD Flat Panel Display wi... 1 LED Edgelight System 1 LED-Backlit LCD Flat Panel Display 1

Related categories: Gaming Accessories Monitors

Additional filters: Flat Panel LCD Monitors x Clear

Results 1-11 of 11 for 4k monitor



UltraSharp 27 4K Monitor: U2718Q



24 Ultra HD 4K Monitor - P2415Q



43 Ultra HD 4K Multi Client Monitor - P4317Q



UltraSharp 32 4K USB-C Monitor: U3219Q



UltraSharp 27 USB-C Monitor: U2719DC



UltraSharp 27 Monitor - U2719D

| Market Value  | \$1,099.99 | Market Value  | \$758.99 | Market Value  | \$599.99 |
|---------------|------------|---------------|----------|---------------|----------|
| Total Savings | \$378.40   | Total Savings | \$231.88 | Total Savings | \$204.00 |
| Price         | \$721.59   | Price         | \$527.11 | Price         | \$395.99 |

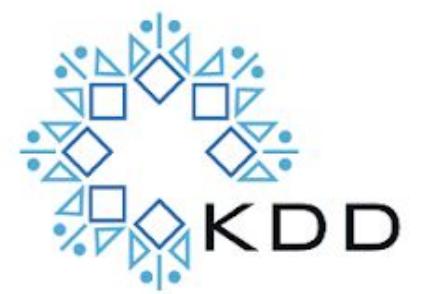
| Market Value  | \$758.99 | Market Value  | \$599.99 |
|---------------|----------|---------------|----------|
| Total Savings | \$231.88 | Total Savings | \$204.00 |
| Price         | \$527.11 | Price         | \$395.99 |

| Market Value  | \$599.99 |
|---------------|----------|
| Total Savings | \$204.00 |
| Price         | \$395.99 |

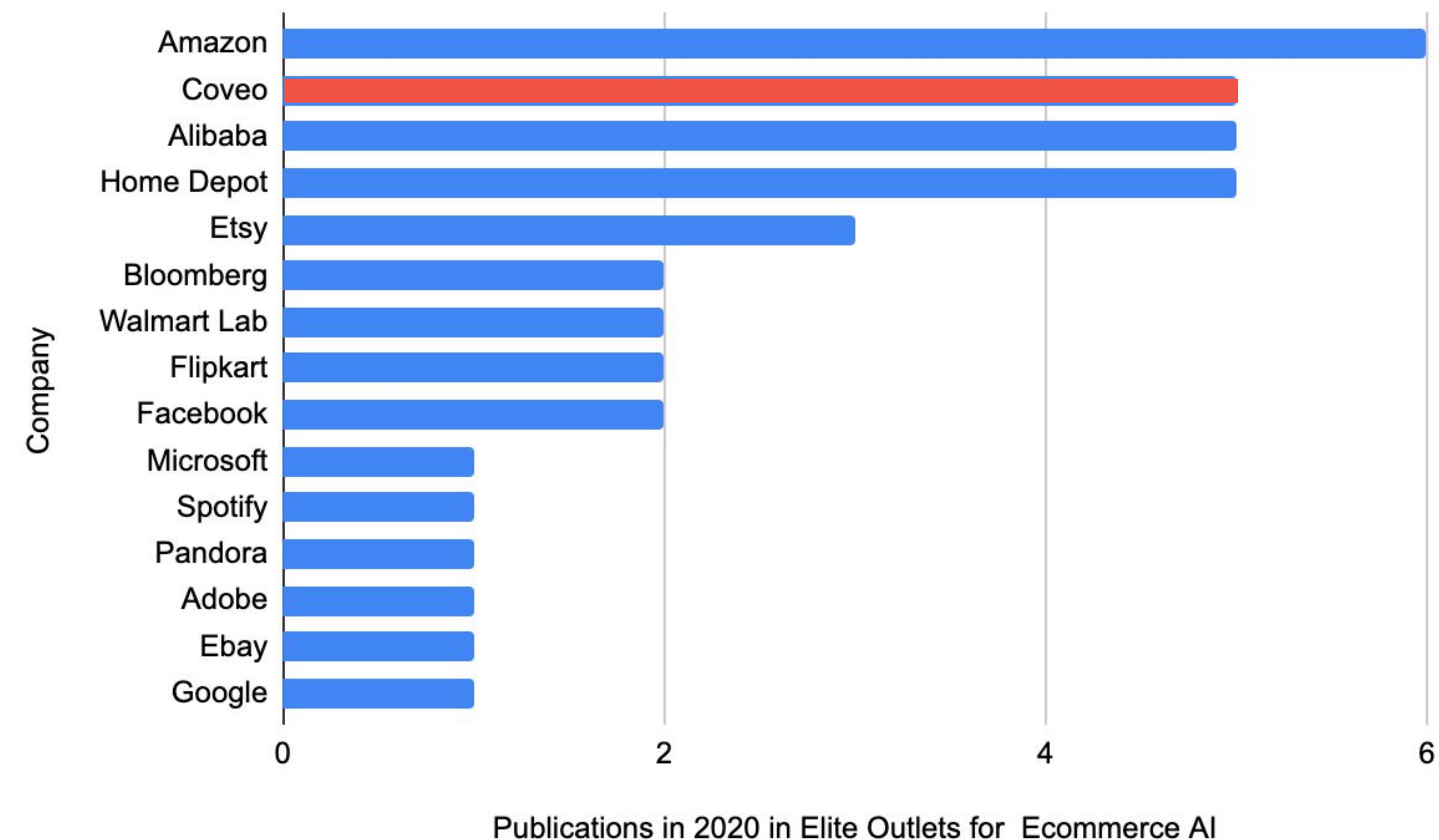
## Coveo's typical integration for Commerce:

- Search
- Query suggestions
- Recommendations
- Category listing

# Publishing in Leading Outlets for AI & NLP

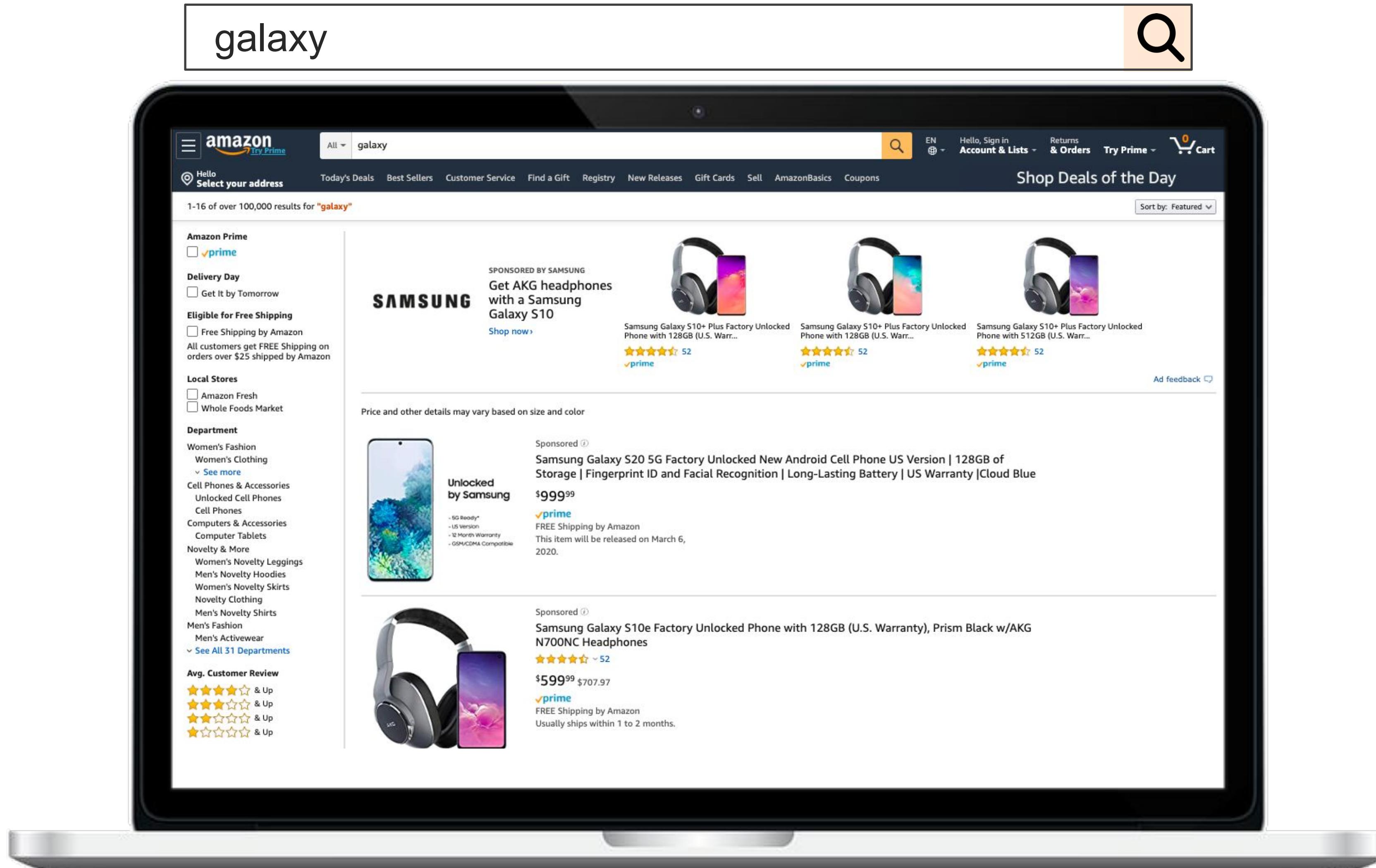


# Leading AI research in Ecommerce

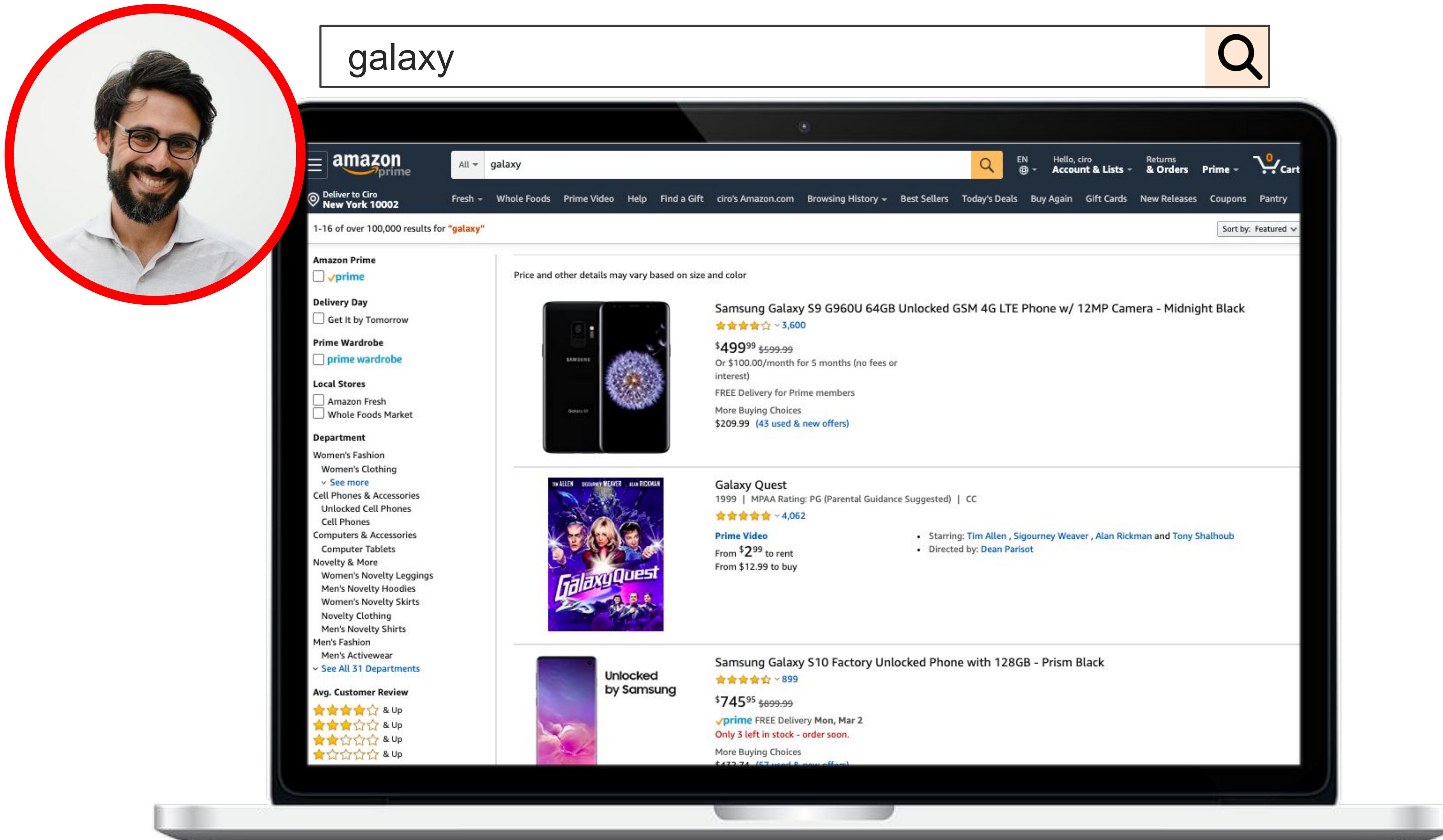


# From personalization to in-session personalization

# Personalization at the time of Amazon



# Personalization at the time of Amazon



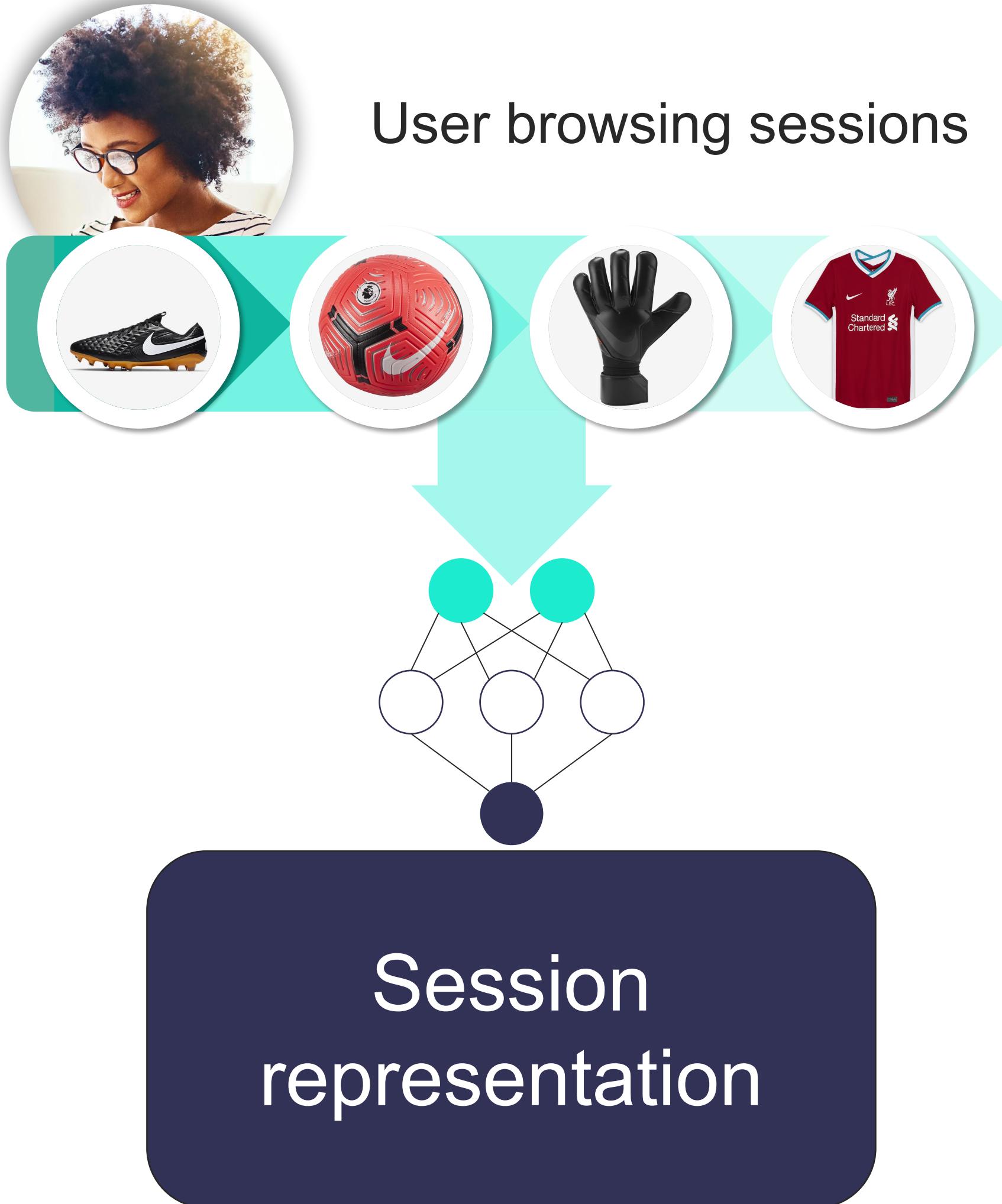
# Goals

- Personalization in digital experience has a direct impact on conversion rate and average cart value.
- Many e-commerce players need personalized search and recommendations but cannot build them themselves.
- Our goal is to develop effective personalization strategies for mid-to-large market segment.

# Constraints

- Most shops don't generate massive amount of data.
- Shoppers are almost never logged-in and don't come back often.
- Personalization needs to happen as early as possible and with as little data as possible.
- The session becomes the natural boundary for our ML models to work.

# A research program for session-based inference



## Use cases

Search

C&LING  
2020

NAACL 2021

Recommendations

RecSys

Query suggestion

ACL

Causal Attribution

SIGIR 2020

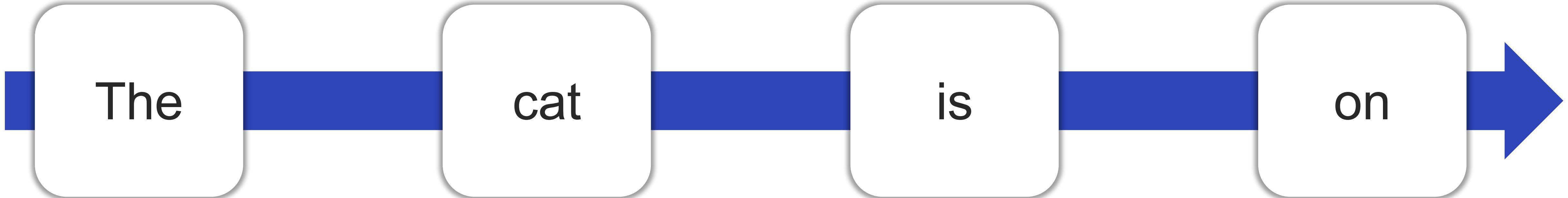
Intent Detection

KDD

# (Product) space, the final frontier

# Product embeddings

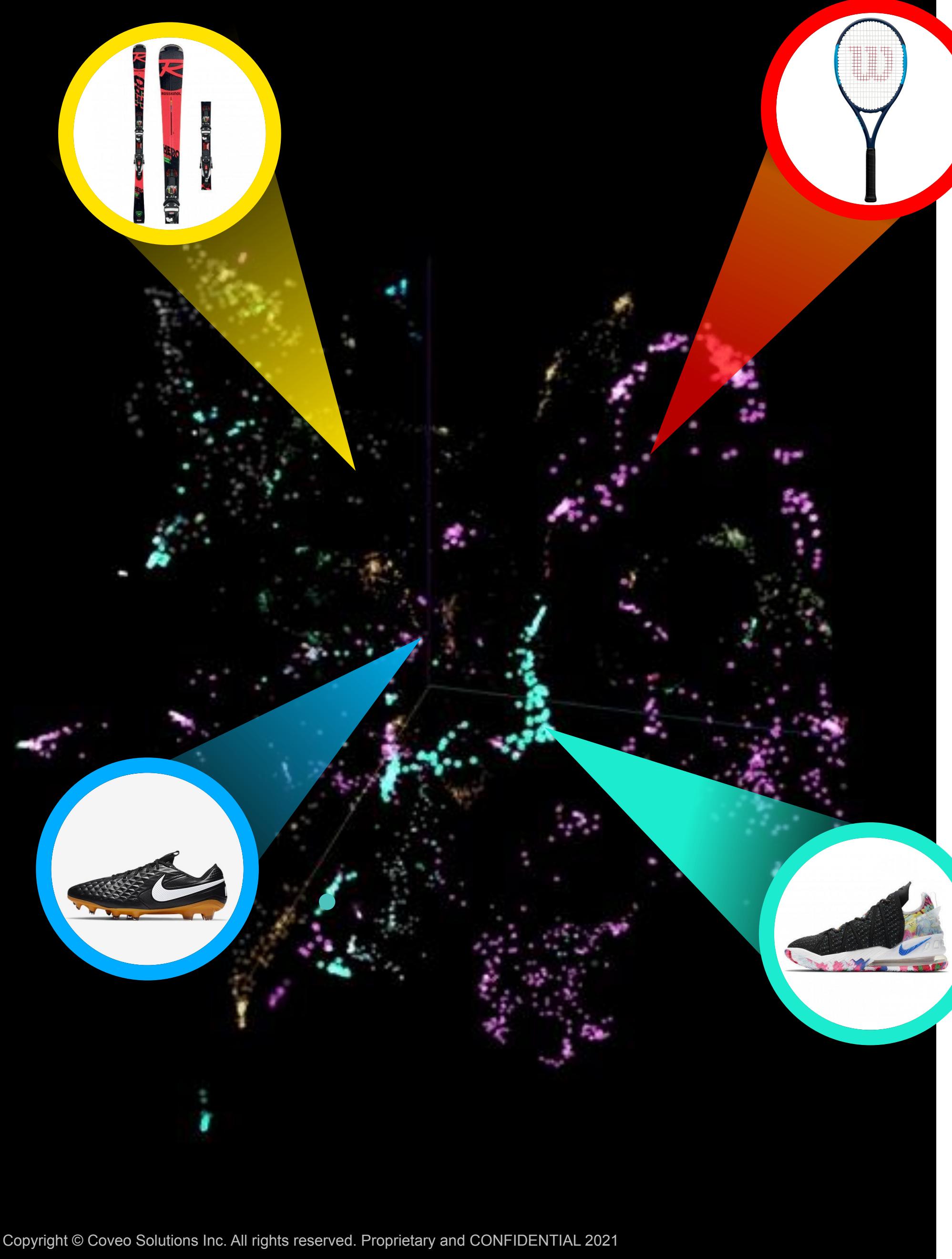
word2vec



prod2vec



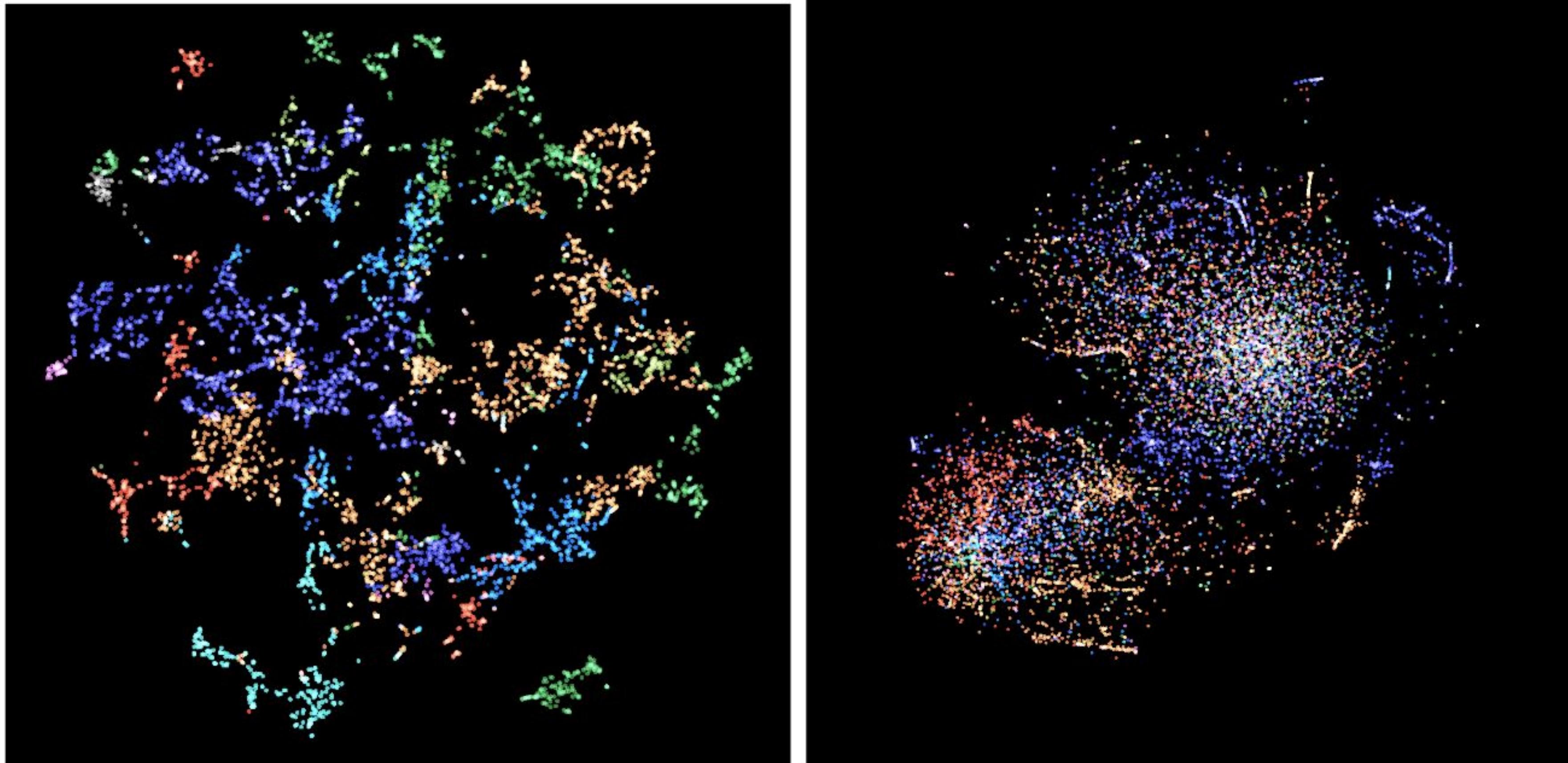
# The product space



- Intent is represented as the products users interact with within a session.
- Products are represented as a multi-dimensional vector space: similar products that are “closer” in the space.
- Building such a space can be done in a purely unsupervised manner.

# Practical Tip #1: tune your hyperparameters

- Embeddings from a sport apparel e-commerce website (colors represent sport activities).



# Practical Tip #2: pay attention to low-count items

- Due to the long-tail of interactions, not all product vectors have the same quality...

*kNN of a popular product*



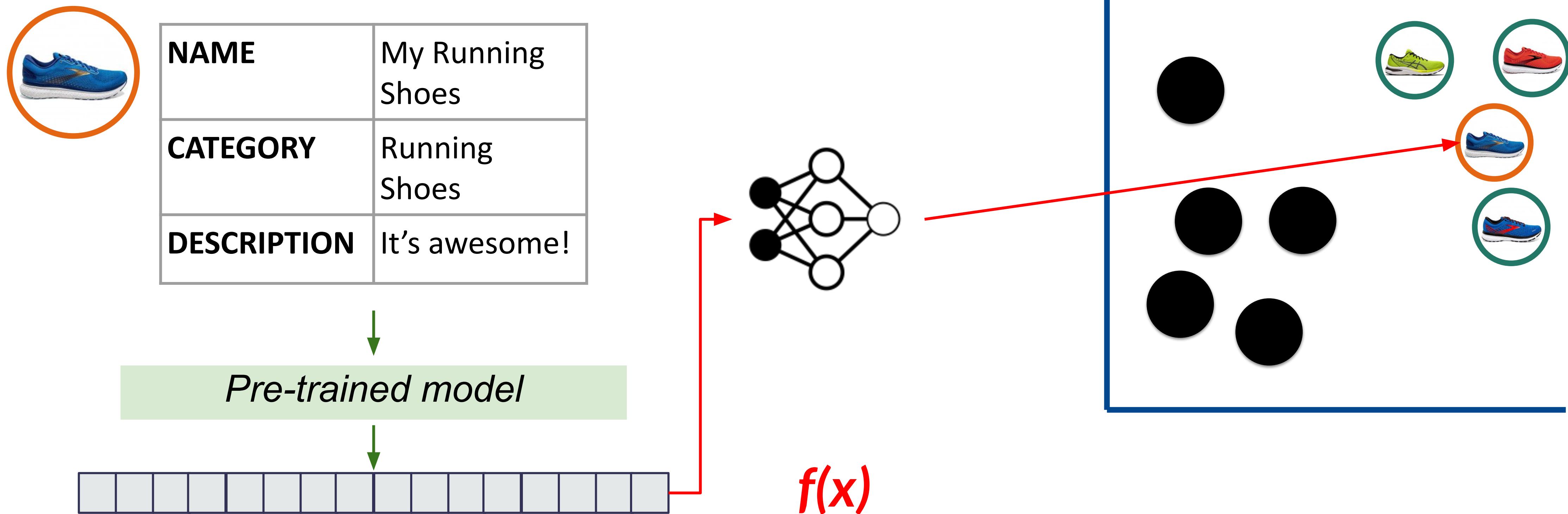
*kNN of a rare product*



*Plus, new products have no interactions!*

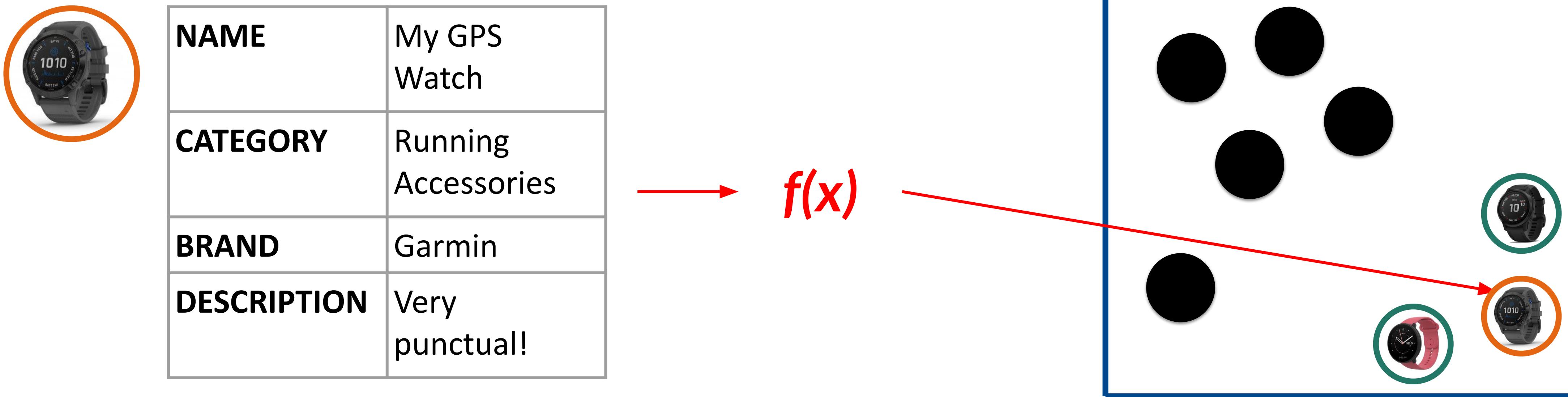
# “Faking” high quality vectors through content

- First, we learn a mapping between meta-data and the space by using *popular products only*.



# “Faking” high quality vectors through content

- First, we learn a mapping between meta-data and the space by using *popular products only*.
- Then, we apply the mapping to *rare products* and obtain new vectors!



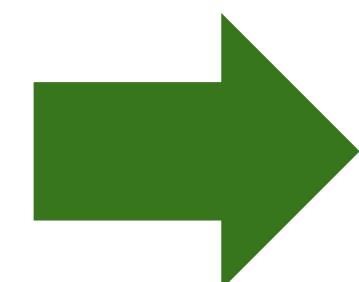
# Practical Tip #3: be *less wrong*

- Considering the use of vectors directly (e.g. kNN for in-session recommendations) or in downstream tasks (e.g. conditional language models), we want our model to be, when not right, *less wrong*.
- When making a mistake on a *rare target item*, we keep track of the *Average Cosine Distance* in the space between the target and prediction

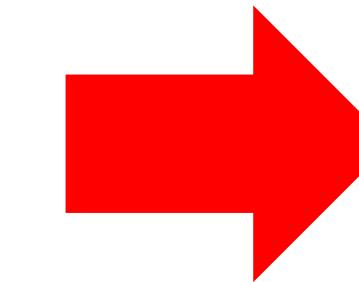
**TARGET**



**PREDICTION**



reasonable



super wrong

# Practical Tip #4: read more!



4906v1 [cs.IR] 20 Jul 2020

## Fantastic Embeddings and How to Align Them: Zero-Shot Inference in a Multi-Shop Scenario

Federico Bianchi\*  
Bocconi University  
Milano, Italy  
f.bianchi@unibocconi.it

Jacopo Tagliabue\*†  
Coveo Labs  
New York, NY  
jtagliabue@coveo.com

Bingqing Yu\*  
Coveo  
Montreal, Canada  
cyu2@coveo.com

Luca Bigon‡  
Coveo  
Montreal, Canada  
lbigon@coveo.com

Ciro Greco§  
Coveo Labs  
New York, NY  
cgreco@coveo.com

### ABSTRACT

This paper addresses the challenge of leveraging multiple embedding spaces for multi-shop personalization, proving that zero-shot inference is possible by transferring shopping intent from one website to another without manual intervention. We detail a machine learning pipeline to train and optimize embeddings *within shops* first, and support the quantitative findings with additional qualitative insights. We then turn to the harder task of using learned embeddings *across shops*: if products from different shops live in the same vector space, user intent - as represented by regions in this space - can then be transferred in a zero-shot fashion across websites. We propose and benchmark unsupervised and supervised methods to “travel” between embedding spaces, each with its own assumptions on data quantity and quality. We show that zero-shot personalization is indeed possible at scale by testing the shared embedding space with two downstream tasks, event prediction and type-ahead suggestions. Finally, we curate a cross-shop

### ACM Reference Format:

Federico Bianchi, Jacopo Tagliabue, Bingqing Yu, Luca Bigon, and Ciro Greco. 2020. Fantastic Embeddings and How to Align Them: Zero-Shot Inference in a Multi-Shop Scenario. In *Proceedings of ACM SIGIR Workshop on eCommerce (SIGIR eCom'20)*. ACM, New York, NY, USA, 11 pages.

### 1 INTRODUCTION

Inspired by the similarity between words in sentences and products in browsing sessions, recent work in recommender systems re-adapted the NLP CBOW model [20] to create *product embeddings* [17], i.e. low-dimensional representations which can be used alone or fed to downstream neural architectures for other machine learning tasks. Product embeddings have been mostly investigated as static entities so far, but, exactly as words [10], products are all but static. Since the creation of embeddings is a stochastic process, training embeddings for similar products in different digital shops



## The Embeddings That Came in From the Cold: Improving Vectors for New and Rare Products with Content-Based Inference

JACOPO TAGLIABUE\*† and BINGQING YU\*, Coveo, United States  
FEDERICO BIANCHI\*, Bocconi University, Italy

Training *product embeddings* in a multi-tenant scenario involves solving the challenges of ever changing catalogs across dozens of deployments, without supervision. In this work, we detail how we deal with new and rare products when building neural representations at scale: we show how to inject product knowledge into behavior-based embeddings to provide the best accuracy with minimal engineering changes in existing infrastructure and without additional manual effort.

CCS Concepts: • Information systems → Clustering; Recommender systems.

Additional Key Words and Phrases: product embeddings, neural networks, cold-start recommendations

### ACM Reference Format:

Jacopo Tagliabue, Bingqing Yu, and Federico Bianchi. 2020. The Embeddings That Came in From the Cold: Improving Vectors for New and Rare Products with Content-Based Inference. In *Fourteenth ACM Conference on Recommender Systems (RecSys '20), September 21–26, 2020, Virtual Event, Brazil*. ACM, New York, NY, USA, 4 pages. <https://doi.org/xx.xxxx/xxxxxx.xxxxxx>

### 1 INTRODUCTION

*Product embeddings* [5, 12] recently became an important part of downstream recommendation [2] and personalization systems [11] given that they capture latent product properties. Embeddings are trained with a CBOW model [8] (*prod2vec*) from behavioral data, resulting in representation of heterogeneous quality – due to the nature of gradient-based training, rare products have lower quality projections [4]. While most research and industry papers focus on producing embeddings at a given time, we address a realistic multi-tenant scenario, in which product spaces evolve over time and across organizations.

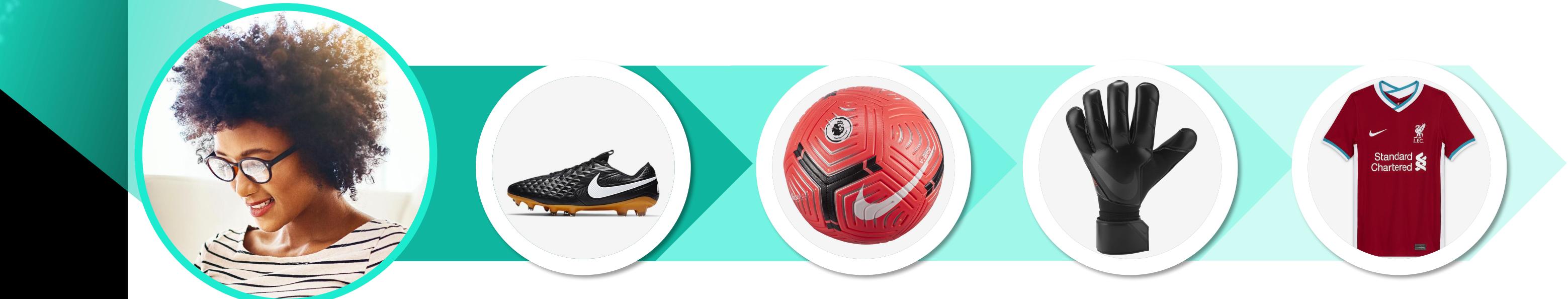
# One representation, to rule them all

# Session representation

- Different users walk in different regions of the space.
- **Session vectors** are functions of the product vectors shoppers interact with:

$$SV = f(p_1-v, p_2-v, \dots p_n-v)$$

[  $f$  can be the (unweighted/weighted) average, or something more complex ]



## Refine by

 For Home For Work

## Product Type

 Accessories

96

## Resolution

 1366 x 768

2

 1600 x 1050

2

 1680 x 1050

5

 1920 x 1080

12

 3840 x 2160

11

## Price

 Under \$500

71

 \$500 to \$1,000

19

 \$1,000 to \$2,500

3

 \$2,500 to \$5,000

2

 \$5,000 and more

1

## Display Size

 65 inches and more

2

 49 to 60 inches

6

 32 to 48 inches

6

 24 to 32 inches

22

Results 1-15 of 96 for 4k monitor



UltraSharp 27 4K Monitor: U2718Q



24 Ultra HD 4K Monitor - P2415Q



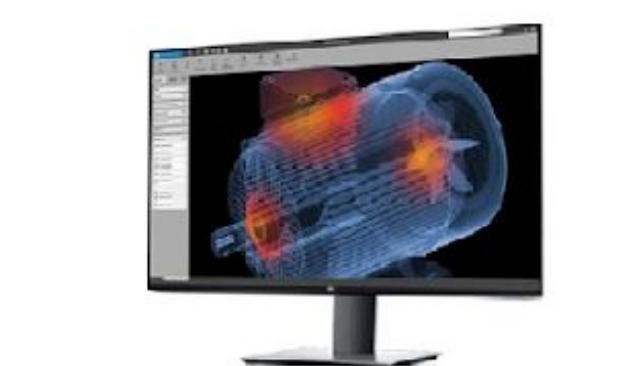
43 Ultra HD 4K Multi Client Monitor - P4317Q



UltraSharp 27 4K HDR Monitor: UP2718Q



UltraSharp 32 Ultra HD 4K Monitor with PremierColor - UP3216Q



UltraSharp 32 4K USB-C Monitor: U3219Q



UltraSharp 27 USB-C Monitor: U2719DC



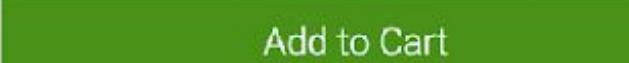
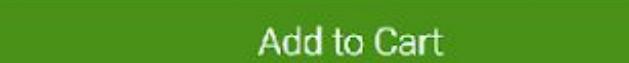
UltraSharp 27 Monitor: U2719D



55 4K Conference Room Monitor: C5519Q

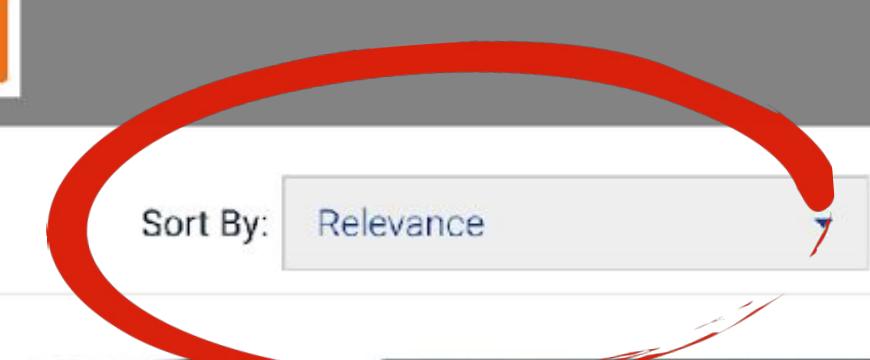


75 4K Interactive Touch Monitor: C7520QT



Sort By:

Relevance



Sort By: Price Low To High

## Refine by

 For Home For Work

## Product Type

 Accessories

102

 Desktops & Laptops

6

**Kensington SD1500 USB-C Mobile Dock - 4K HDMI or HD VGA - Windows/Chrome/Mac - Docking station - USB - VGA, HDMI - GigE**

## Resolution

 1366 x 768

2

 1600 x 1050

2

 1680 x 1050

5

 1920 x 1080

12

 3840 x 2160

11



## Price

 Under \$500

41

 \$500 to \$1,000

22

 \$1,000 to \$2,500

36

 \$2,500 to \$5,000

8

 \$5,000 and more

1

**StarTech.com Thunderbolt 3 Dual Monitor HDMI Adapter - 4K30 - Windows only - external video adapter**

## Display Size

 65 inches and more

17

 49 to 60 inches

12

 32 to 48 inches

2

**Targus USB-C Multiport Video Adapter - External video adapter - USB-C - HDMI - black****StarTech.com Thunderbolt 3 to Dual DisplayPort Adapter - 4K 60Hz - Mac and Windows Compatible - External video adapter - Thunderbolt 3 - 2 x DisplayPort - Silver****Kensington SD1600P USB-C Mobile 4K Dock with Pass-Through Charging - Docking station - USB-C - VGA, HDMI - GigE****Belkin Multiport to HDMI Digital AV Adapter 8 ft - 4K Support****Docking Station - USB 3.0 (D3100)****Targus Universal 4k Docking Station - Docking station - (USB) - GigE - US****Add to Cart****Add to Cart****Add to Cart****Add to Cart****Add to Cart**

## Refine by

- For Home
- For Work

### Product Type

- Accessories

96

**UltraSharp 27 4K Monitor: U2718Q****24 Ultra HD 4K Monitor - P2415Q****43 Ultra HD 4K Multi Client Monitor - P4317Q****UltraSharp 27 4K HDR Monitor: UP2718Q****UltraSharp 32 Ultra HD 4K Monitor with PremierColor - UP3216Q**

### Resolution

- 1366 x 768

2

|               |          |
|---------------|----------|
| Market Value  | \$739.99 |
| Total Savings | \$238.40 |
| Price         | \$501.59 |

|               |          |
|---------------|----------|
| Market Value  | \$549.99 |
| Total Savings | \$171.60 |
| Price         | \$378.39 |

|               |            |
|---------------|------------|
| Market Value  | \$1,172.88 |
| Total Savings | \$360.75   |
| Price         | \$812.13   |

|               |            |
|---------------|------------|
| Market Value  | \$1,099.99 |
| Total Savings | \$697.60   |
| Price         | \$1,302.39 |

|               |            |
|---------------|------------|
| Market Value  | \$1,799.99 |
| Total Savings | \$612.00   |
| Price         | \$1,187.99 |

### Price

- Under \$500

71

**UltraSharp 32 4K USB-C Monitor: U3219Q****UltraSharp 27 USB-C Monitor: U2719DC****UltraSharp 27 Monitor: U2719D****55 4K Conference Room Monitor: C5519Q****75 4K Interactive Touch Monitor: C7520QT**

- \$5,000 and more

1

|               |            |
|---------------|------------|
| Market Value  | \$1,099.99 |
| Total Savings | \$378.40   |
| Price         | \$721.59   |

|               |          |
|---------------|----------|
| Market Value  | \$758.99 |
| Total Savings | \$231.88 |
| Price         | \$527.11 |

|               |          |
|---------------|----------|
| Market Value  | \$599.99 |
| Total Savings | \$204.00 |
| Price         | \$395.99 |

|               |            |
|---------------|------------|
| Market Value  | \$1,149.99 |
| Total Savings | \$402.00   |
| Price         | \$747.99   |

|       |            |
|-------|------------|
| Price | \$5,999.99 |
|-------|------------|

### Display Size

- 65 inches and more

2

|             |
|-------------|
| Add to Cart |
|-------------|

|             |
|-------------|
| Add to Cart |
|-------------|

|             |
|-------------|
| Add to Cart |
|-------------|

|             |
|-------------|
| Add to Cart |
|-------------|

|             |
|-------------|
| Add to Cart |
|-------------|

- 49 to 60 inches

6



- 32 to 48 inches

6



- 24 to 32 inches

22



4k monitor



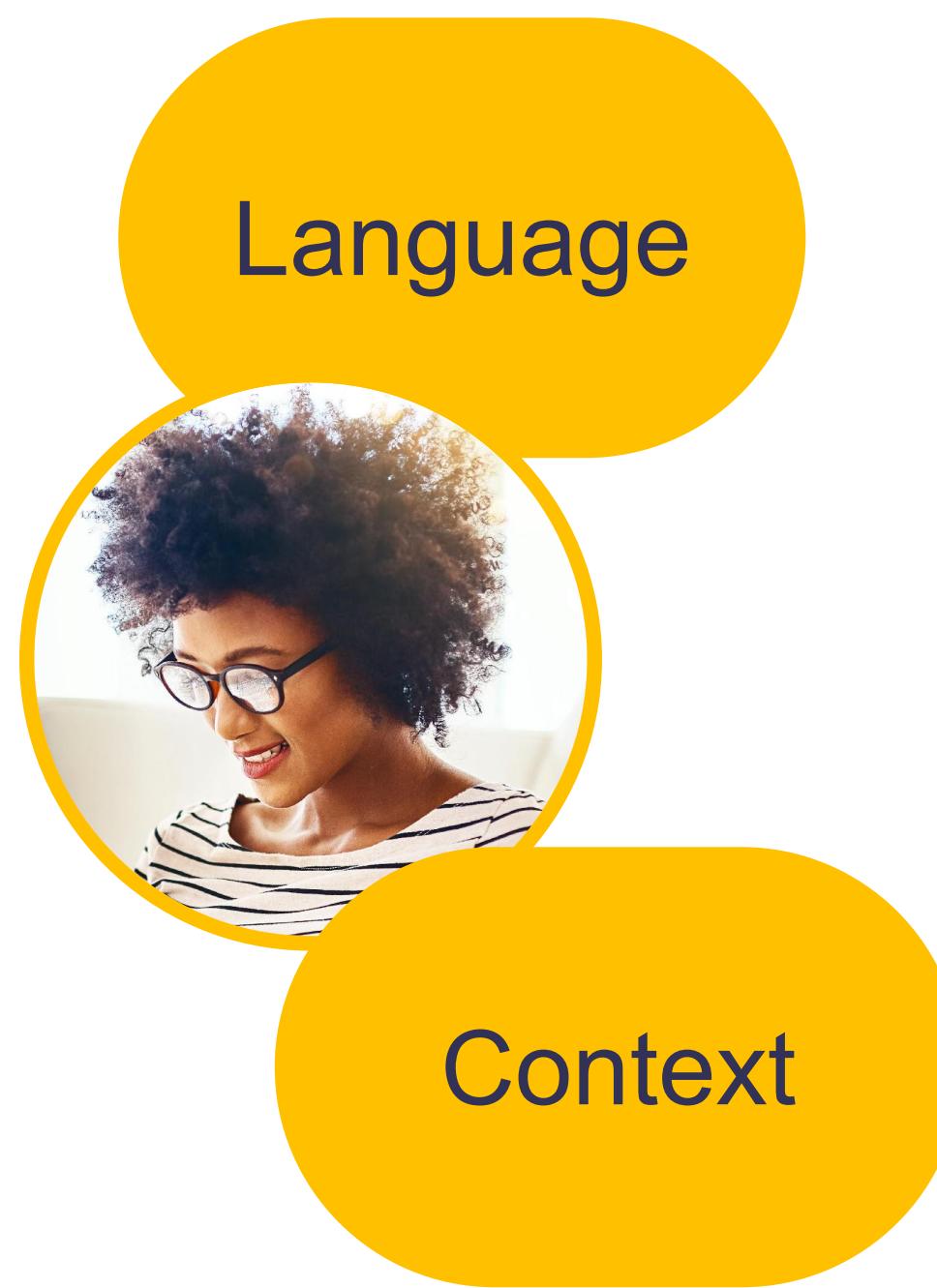
Search

Sort By:

Relevance



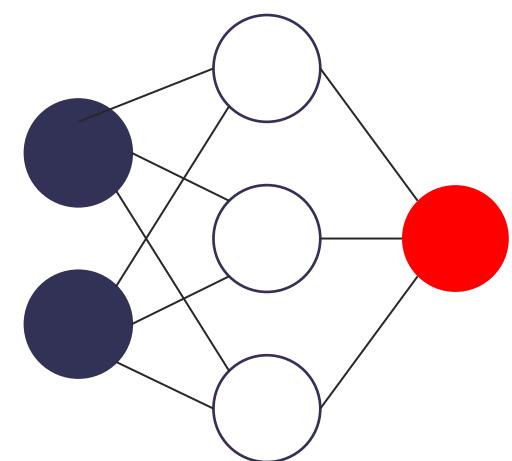
# In-session query scoping



4k monitor



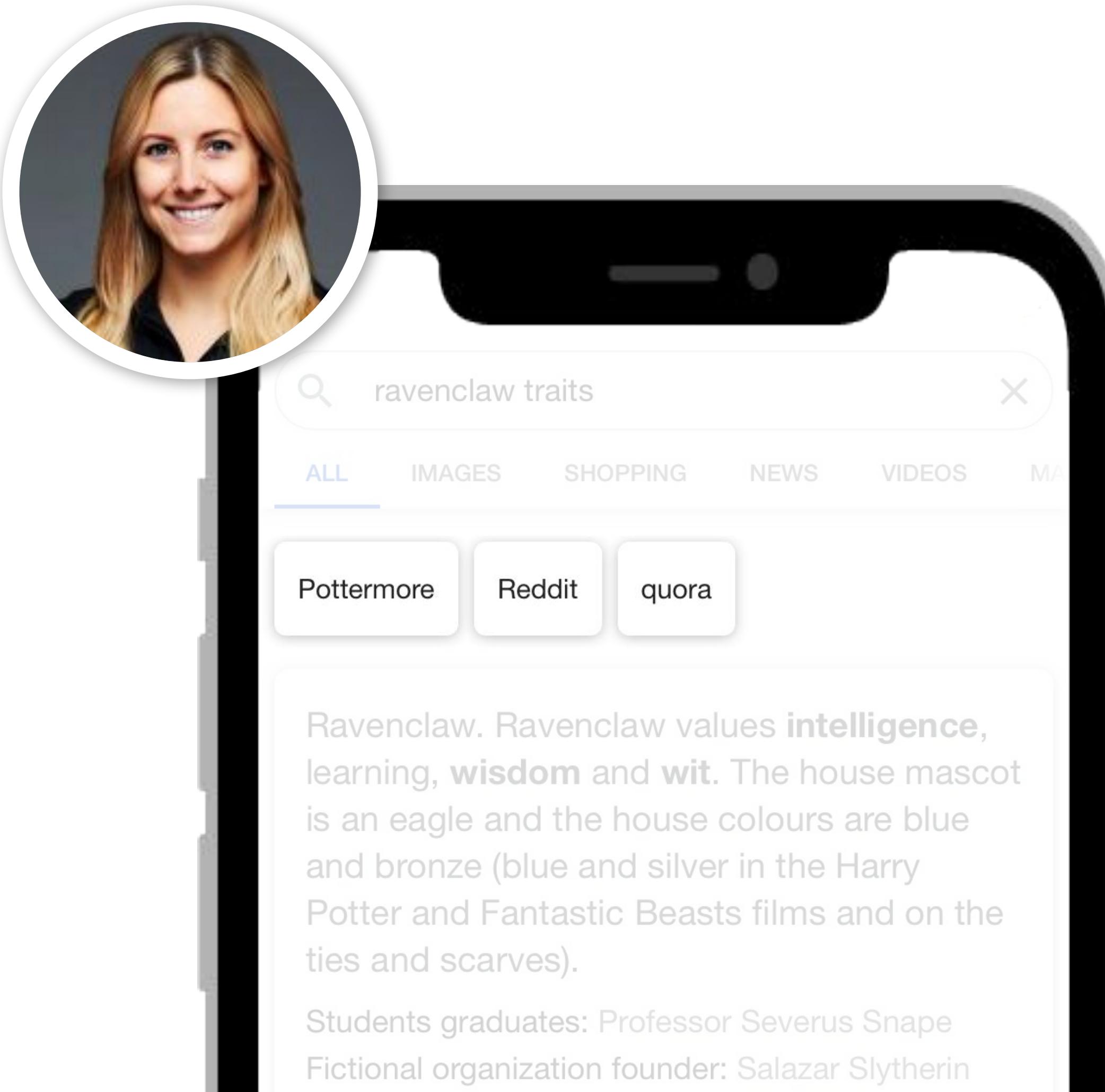
Query vector



Session vector

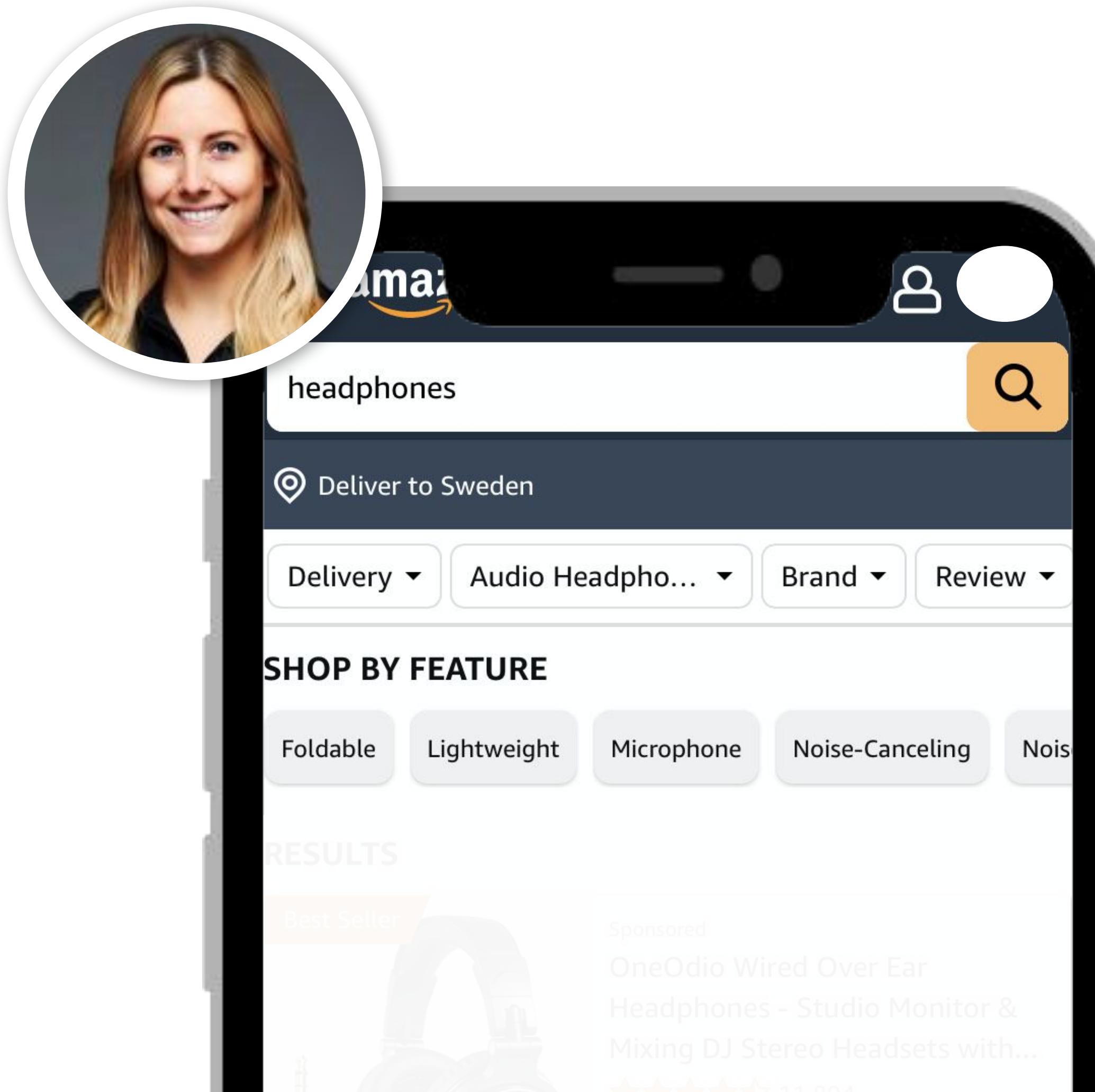
Monitors

# Help product discovery with “discovery tags”



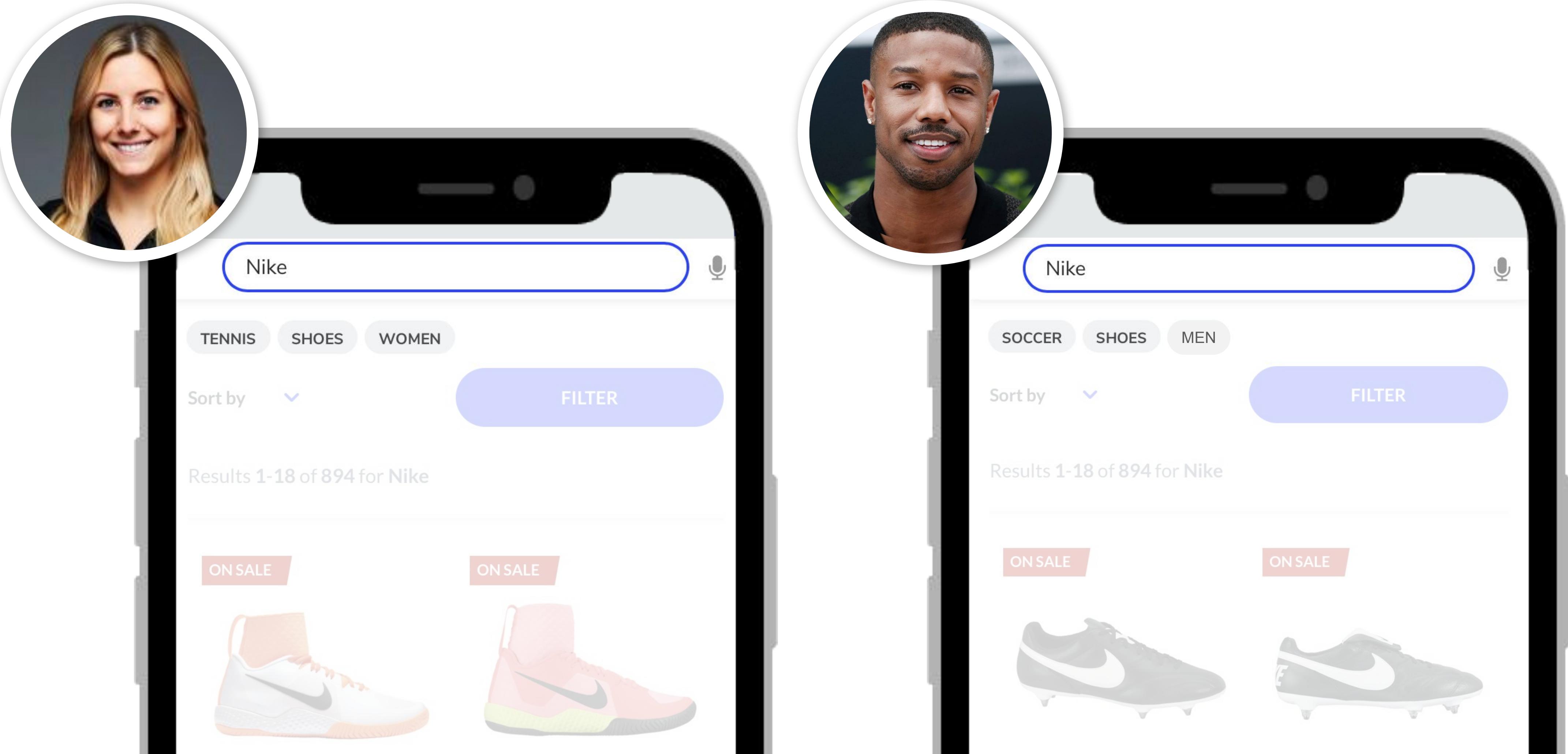
Google introduced mobile-friendly query refinements in 2020...

# Help product discovery with “discovery tags”

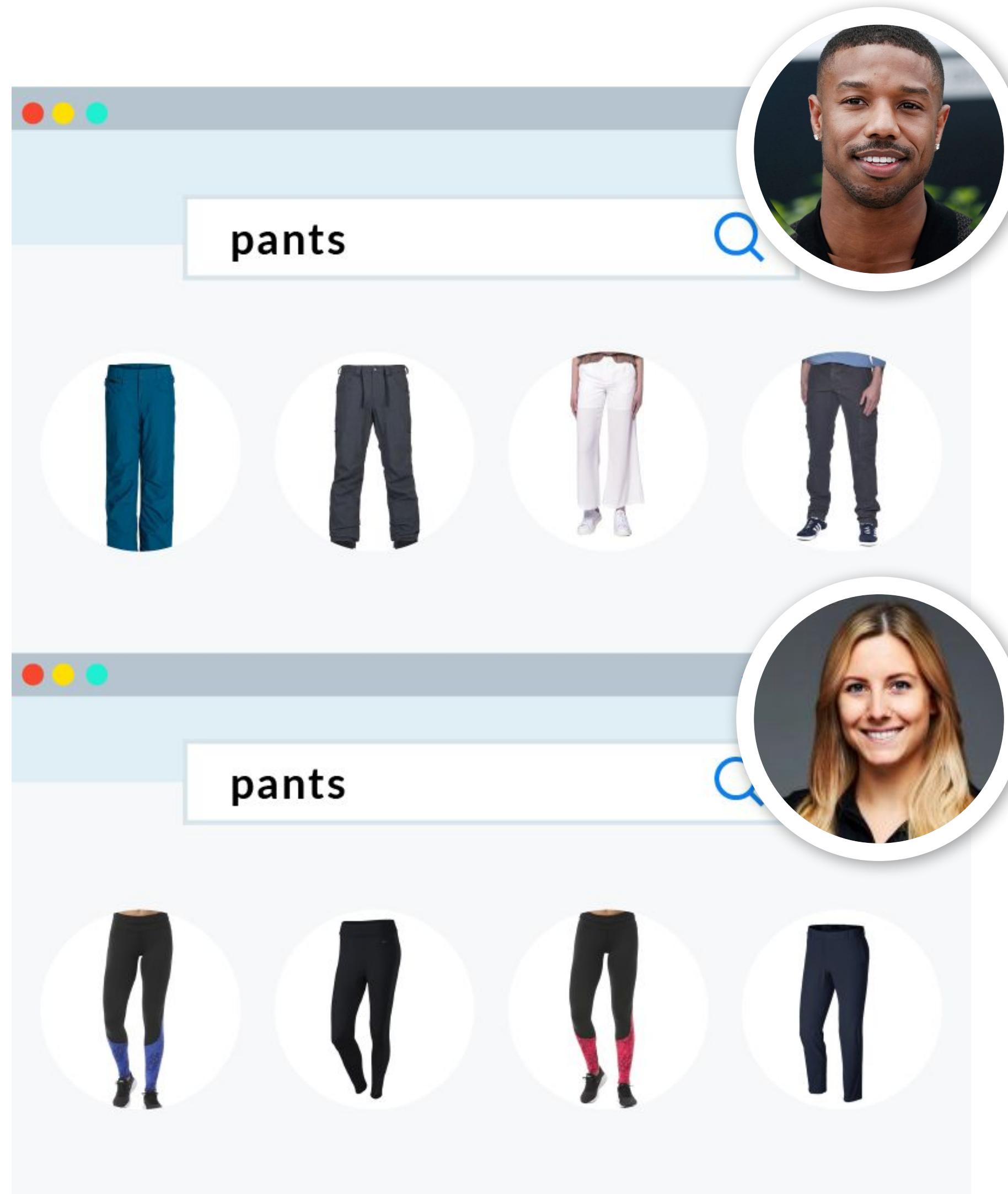


...and Amazon  
followed immediately  
afterwards!

# Help product discovery with “discovery tags”



# Dynamically re-rank search results



**Clothes in Space: Real-time personalization in less than 100 lines of code**



Jacopo Tagliabue December 11, 2019

TOPICS

- / ai
- / ecommerce
- / machine learning
- / natural language processing
- / personalization
- / product embeddings

## Search is About to Get Personal

*"WHO WANTS TO BE NORMAL WHEN YOU CAN BE UNIQUE?" H. B. CARTER*

Meet Audrey, Coveo's (part-time) director of marketing and (full-time) long distance runner (notwithstanding the iStockPhoto-like picture, we swear she is a real person). She is browsing products on a sports apparel website, looking at one pair of shoes after the other:

# ...and more!

- There are many other use-cases related to in-session personalization and, generally, a clever use of the product space.
- Have a look at our latest papers to know more, or get in touch with us!

## BERT Goes Shopping: Comparing Distributional Models for Product Representations

Federico Bianchi  
Bocconi University  
Milano, Italy  
[f.bianchi@unibocconi.it](mailto:f.bianchi@unibocconi.it)\*

Bingqing Yu  
Coveo  
Montreal, CA  
[cyu2@coveo.com](mailto:cyu2@coveo.com)

Jacopo Tagliabue  
Coveo AI Labs  
New York, United States  
[jtagliabue@coveo.com](mailto:jtagliabue@coveo.com)

### Abstract

Word embeddings (e.g., word2vec) have been applied successfully to eCommerce products through *prod2vec*. Inspired by the recent performance improvements on several NLP tasks brought by contextualized embeddings, we propose to transfer BERT-like architectures to eCommerce: our model – *ProdBERT* – is trained to generate representations of products through masked session modeling. Through extensive experiments over multiple shops, different tasks, and a range of design choices, we

and Information Retrieval (IR) use cases for eCommerce (Vasile et al., 2018; Bianchi et al., 2020).

As a key improvement over *word2vec*, the NLP community has recently introduced *contextualized representations*, in which a word like *play* would have different embeddings depending on the general topic (e.g. a sentence about *theater* vs *soccer*), where as in *word2vec* the word *play* is going to have only one vector. Transformer-based architectures (Vaswani et al., 2017) in large-scale models – such as BERT (Devlin et al., 2019) – achieved

## Shopping in the Multiverse: A Counterfactual Approach to In-Session Attribution

Jacopo Tagliabue\*†  
[jtagliabue@coveo.com](mailto:jtagliabue@coveo.com)  
Coveo Labs  
New York, USA

Bingqing Yu\*  
[cyu2@coveo.com](mailto:cyu2@coveo.com)  
Coveo  
Montreal, Canada

### ABSTRACT

We tackle the challenge of in-session attribution for on-site search engines in eCommerce. We phrase the problem as a causal counterfactual inference, and contrast the approach with rule-based systems from industry settings and prediction models from the multi-touch attribution literature. We approach counterfactuals in analogy with treatments in formal semantics, explicitly modeling possible outcomes through alternative shopper timelines; in particular, we propose to learn a generative browsing model over a target shop, leveraging the latent space induced by *prod2vec* embeddings; we show how natural language queries can be effectively represented in the same space and how “search intervention” can be performed to assess causal contribution. Finally, we validate the methodology on a synthetic dataset, mimicking important patterns emerged in customer interviews and qualitative analysis, and we

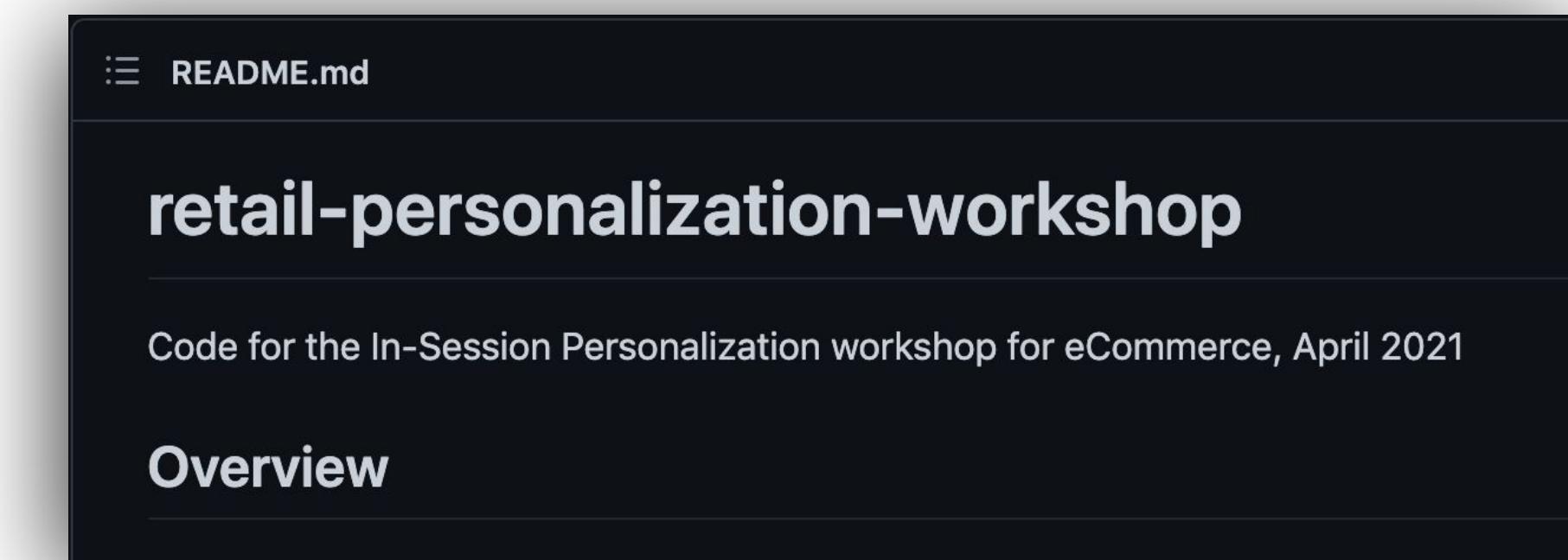
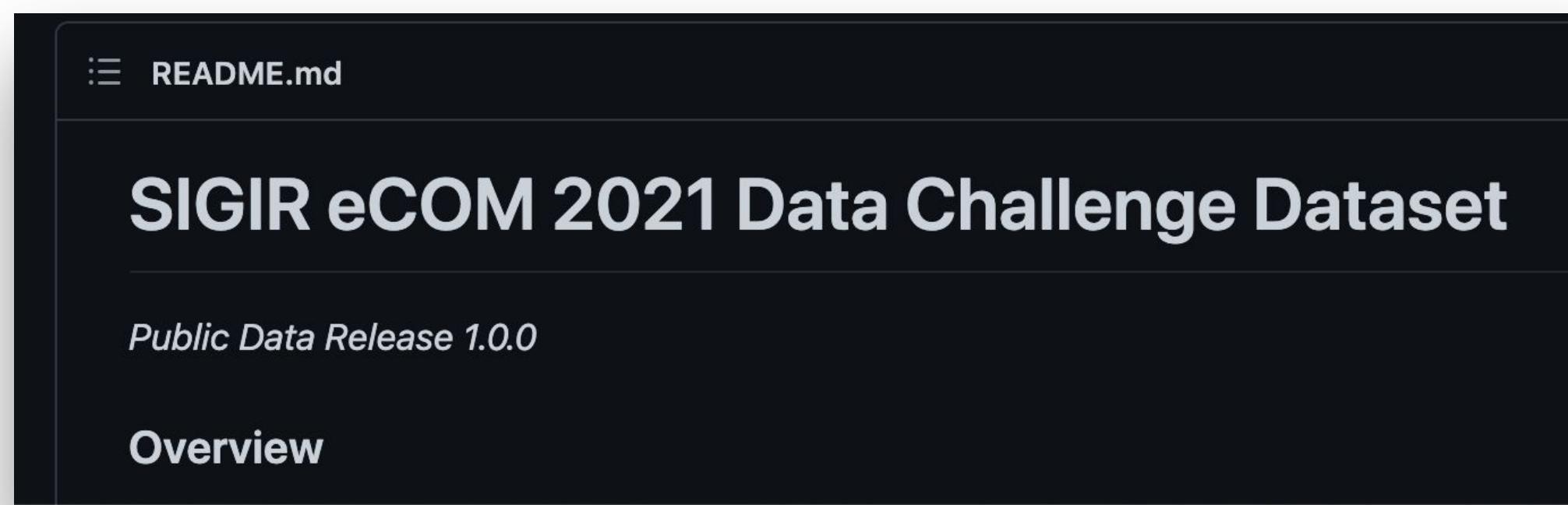
categories: socks, t-shirts, headbands... finally, Simon finds a pair of shorts that he really likes, adds them to cart and completes the transaction. Later that day, Alice, *Balls&Things* director of digital marketing, is looking at *Google Analytics* conversion dashboard: Simon’s session is not just a win (i.e. *conversion*), it is a win for *on-site search*.

*Is it, though?* In this case, it seems pretty obvious that the search behavior and the conversion event are only mildly related, but some other cases are subtler: did our search engine lead Simon to the purchase, or would he have bought shorts anyway? *On-site attribution* is the task of determining the value of each on-site customer touchpoint, such as on-site search, that leads to a conversion. Addressing the challenge in a principled way is important to many stakeholders: Alice, who needs to allocate budget depending on conversion signals; John, *Balls&Things* CTO, who needs to measure

# Hands-on coding

# Open source and open data workshop

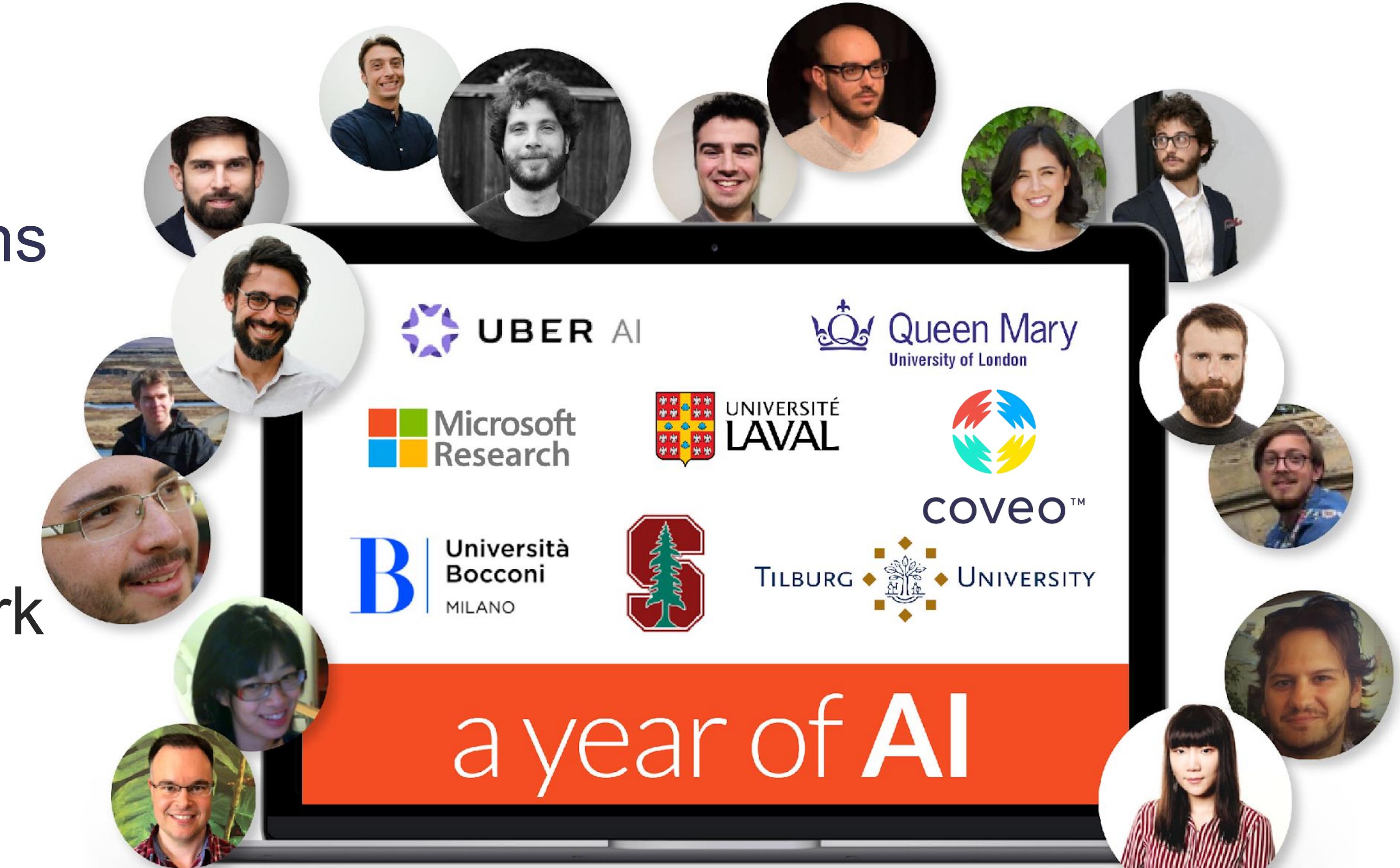
- Let's see some of these ideas at work with *real code AND real data!*
- Code is shared freely under MIT license on [Github](#)
- If you want to start playing along with some real data, you can download the SIGIR Data Challenge dataset from Coveo's [Github!](#)



# What's next for us?

# What's next?

- In 2021, we will be working on query semantics, small-data recommendations and our experimental platform.
- We have a small-but-expanding network of collaborations.
- See something you like?  
**Let's join forces!**



See you, space cowboys.



coveo™