

Real-world Applications of LLMs for eCommerce

Ching-Wei Chen
VP of Engineering, Crossing Minds

Presented at ISIR-eCom @ WSDM 2024
Mar 8th, 2024, Mérida, Mexico

Who I am

Since 2022: VP of Engineering at Crossing Minds

Previously: Director of ML Engineering at Spotify

Experience: Recommender systems, information retrieval, machine learning, signal processing, computer vision

Organizer:

- Spotify Million Playlist Dataset Challenge / RecSys Challenge 2018
- PodRecs @ RecSys 2020, 2021: Workshop on Podcast Recommendations



Who is Crossing Minds?

- An **AI Personalization Platform** founded in 2017 by Alexandre Robicquet, Emile Contal, and Sebastian Thrun
- We provide a highly scalable **Recommendations API**, along with **AI-based Data Enrichment** and **Conversational Search**.
- We serve over 30 customers in the media, financial, education, and **eCommerce** industries.
- Within eCommerce, we work with a number of **brands**, **retailers**, and **marketplaces**.

CHANEL



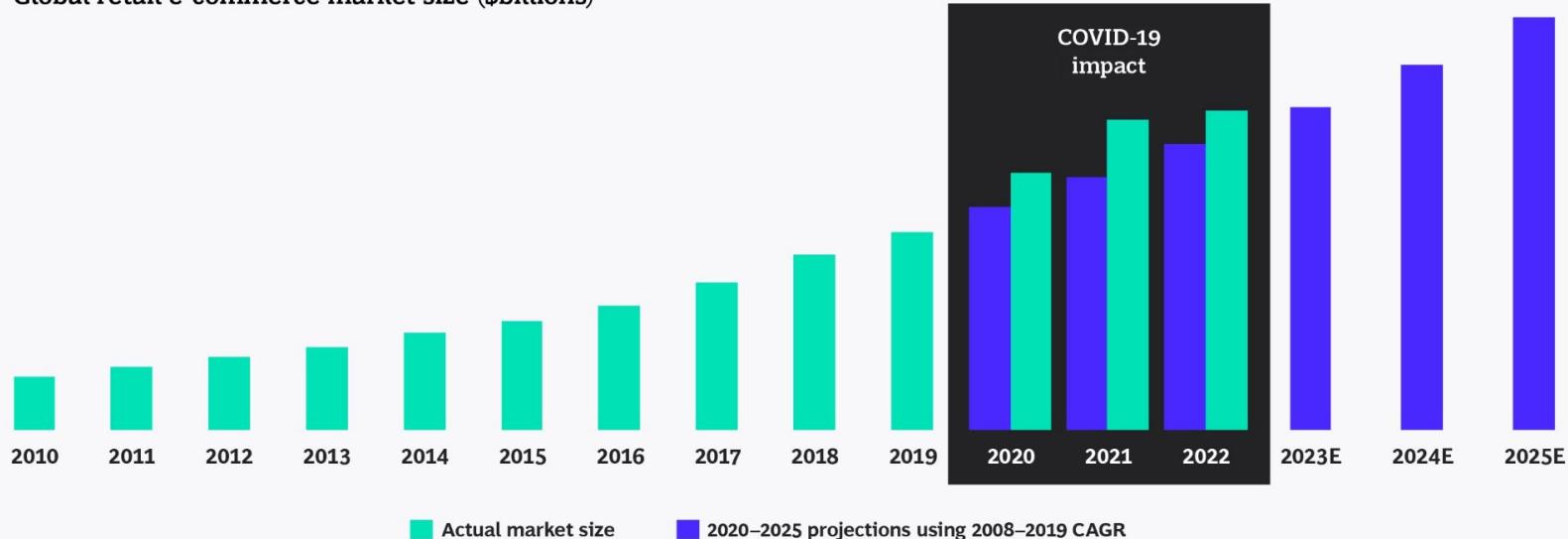
Bonhams

*Rails***Flink'****KOHL'S** **shopify** **DRINKS** **dutchie**

Trends in eCommerce

E-commerce acceleration slowed in 2021–2022 but is returning to its long-term trend line

Global retail e-commerce market size (\$billions)



Source: Euromonitor.

Note: Growth is worldwide and includes all categories and all retail and CPG e-commerce distribution types (pure, specialty, and grocery players).

E-commerce to lead global retail sales growth by 2027

Share of retail sales (%)

GLOBAL RETAIL
(\$TRILLIONS)

8

11

4

15

CAGR
2022–2027

9%

4%



Sources: BCG's Worldwide E-Commerce Survey 2023; Foxintelligence; Edge by Ascential.

Note: E-commerce includes both retail and CPG companies.

Mobile Commerce

- In 2023, Mobile Commerce accounted for 42% of all eCommerce sales

Retail Mcommerce Sales in the US, 2019-2025

billions, % change, and % of total retail ecommerce sales

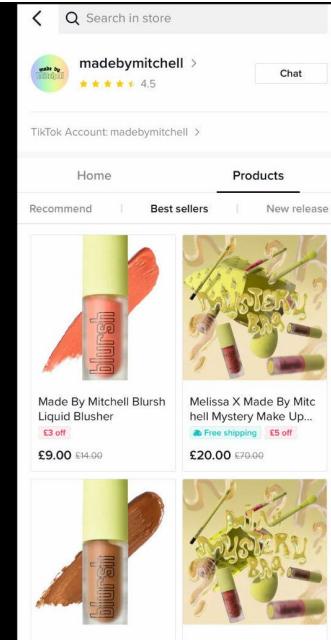
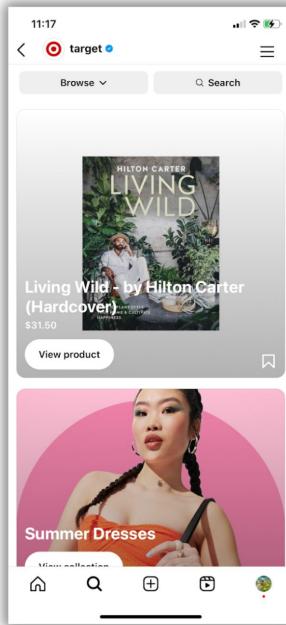


Note: includes products or services ordered using the internet via mobile devices, regardless of the method of payment or fulfillment; includes sales on tablets; excludes travel and event tickets, payments such as bill pay, taxes, or money transfers, food services and drinking place sales, gambling, and other vice goods sales

Source: eMarketer, May 2021

Social Commerce

- Social Commerce in the US accounted for \$56.2B in 2023, and is projected to grow to \$79.6B by 2025.



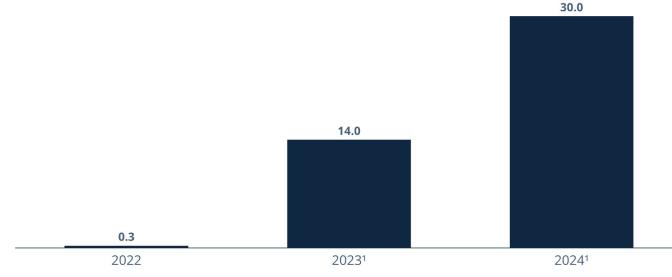
Shifts in distribution

- **Shopify** and other integrated eCommerce platforms have enabled **sellers of all sizes** to enjoy the same tech stack enjoyed by only the larger online stores decades ago. The number of merchants on Shopify has grown from 1M in 2019, to **over 2.3M in 2024**.
- Temu, Shein, Aliexpress and other **manufacturer-to-consumer (M2C) marketplaces** have had a meteoric rise in the last few years, capitalizing on budget-conscious shoppers in a high-inflation financial climate.



Source: [Store Leads](#)

GROSS MERCHANDISE VOLUME (GMV) ON TEMU.COM, 2022-2024
in billion US\$



Notes:(1) Forecast.
Sources: ECDB.

ECDB

Source: [ECDB](#)

Challenges in eCommerce

Challenges faced by eCommerce businesses

Content

Rich product metadata is **crucial** for discoverability both on site and off, but metadata is costly to manually create, and quality and coverage can vary wildly.

UX

Modern shoppers are beginning to shift from “search engine” mindsets to personalized feeds, and conversational interfaces.

Enter the LLM



- Large Language Models, or **LLMs** were built upon prior research in Deep Learning, particularly the Transformer architecture.
- ChatGPT by OpenAI was the most successful application of LLMs that ushered in the current era of Generative AI in mainstream culture.
- What exactly can LLMs do?

LLMs can talk

FL

Hi, I'm writing an article about you and all of the cool things you can do. 
Could you say hello to my readers and give us a quick self-introduction? I'd like it to be friendly and casual, include a silly joke, and end with a note about how great the rest of my article is.



Hello readers! My name is Assistant, and I'm a large language model trained by OpenAI. I'm here to help answer your questions and provide you with information on a wide variety of topics. I'm a bit of a jokester, so here's a quick one for you: Why was the math book sad? Because it had too many problems! But seriously, I'm excited to be featured in your article and I hope you enjoy learning more about me and all the cool things I can do. And trust me, the rest of the article is going to be great!  

LLMs can summarize

Search results for "bryce canyon arches best for family of kids under 3":

- Family Can Travel**
<https://www.familycantravel.com> › bryce-canyon-nati... :
- Ultimate Guide to Bryce Canyon National Park with Kids**
 Jan 25, 2024 — One of the best things to do in Bryce Canyon for kids is the Junior Ranger program (for kids ages 3 and up.) To earn a Bryce Canyon Junior ...
 Hiking Fairyland Loop Trail · Red Canyon Hiking · Kodachrome Basin State Park
- Family Can Travel**
<https://www.familycantravel.com> › arches-national-par... :
- Arches National Park with Kids**
 May 1, 2023 — Playing at the beach under the Sand Dunes Arch is one of the best things to do in Arches for kids.
 1. Delicate Arch Trail · 2. Devil's Garden Loop Trail · 2. Delicate Arch Lookout...
- exploringthroughlife.com**
<https://exploringthroughlife.com> › utah-national-parks... :
- Visiting Utah's National Parks with Young Kids**
 Feb 24, 2021 — One of the most impressive arches in the park is Delicate Arch. If your kids are great hikers, you can hike the 3 miles out to see it up close.
 Can You Road Trip Through... · What To Do At Each Park · Zion National Park With...
- Tripadvisor**
<https://www.tripadvisor.com> ... › Utah Travel Forum :
- Help: Arches or Bryce NP in Mid-February with 2 young kids?**
 Jan 2, 2014 — We didn't plan anything specifically kid related - we mostly just hiked. Arches is a great park for kids because there is a lot to see with ...
 10 answers · Top answer: Just found out that driving from SLC to Arches NP is shorter, 3.5hr, ...
- Full Suitcase**
<https://fullsuitcase.com> › arches-national-park-kids :
- Best Things To Do in Arches National Park with Kids**
 Jan 23, 2024 — If you are visiting Arches with young kids, you can see Delicate Arch in the distance from the Delicate Arch viewpoint along the scenic drive.



Summary of search results:

what's better for a family with kids under 3 and a dog, bryce canyon or Arches National Park...

Generative AI is experimental

Both Bryce Canyon and Arches National Parks are family-friendly. Although both parks prohibit dogs on unpaved trails, Bryce Canyon has two paved trails that allow dogs.

Bryce Canyon has distinctive features like hoodoos, natural bridges, and waterfalls. At the visitor center's interactive exhibits, children can learn about the geology, wildlife, and people of the area. You can bring your dog on the paved sections of both the Rim Trail and the Shared Use Path, both of which are also stroller-friendly.

At Arches, children may enjoy the rock formations, and some say that Arches has more variety than Bryce Canyon. Pets aren't allowed on any trails, though they are allowed at campgrounds, pullouts, roads, and parking lots.

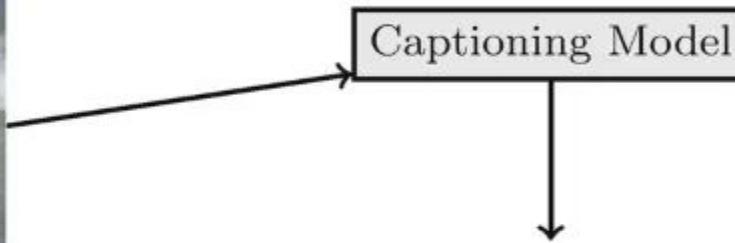
Both parks offer guided tours, ranger-led programs, picnic areas, and restrooms, and both parks require pets to be leashed.

Action Tour Guide
 Feb 17, 2022 — Arches only allow dogs at campgrounds, parking lots, and along established roads, making it a lot harder to experience the park with your dog.

The MOM Trotter
<https://themomtrotter.com> › bryce-canyon-national-pa... :

A Guide To Planning A Trip To Bryce Canyon National ...

LLMs can see



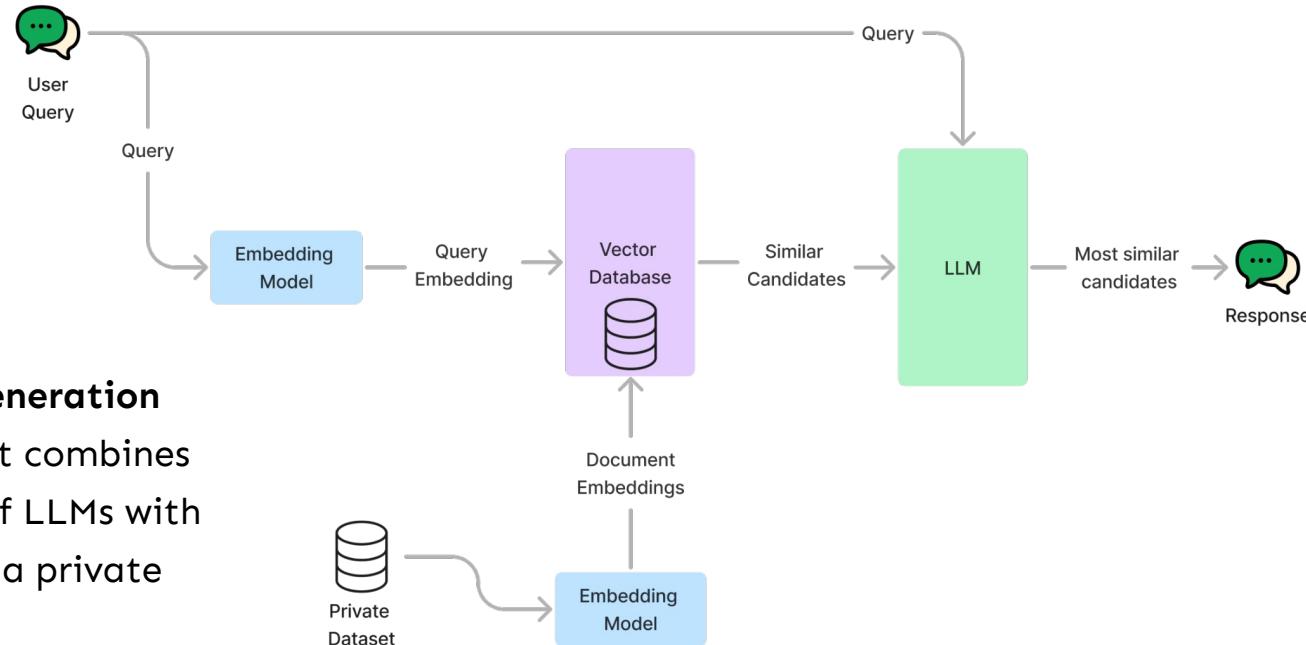
A happy dog is standing in the ocean

What LLMs *aren't* great at

- LLMs are trained on a massive corpus of trillions of words (tokens) from the entire Internet. They encode this knowledge into a model with a much smaller number (still billions) of parameters than the number of inputs.
- Because of this, LLMs cannot reliably recall precise information from the training corpus - in fact, they can "**hallucinate**" and generate outputs that may seem plausible, but are factually incorrect.



Making LLMs better at knowledge-intensive tasks



- **Retrieval Augmented Generation (RAG)** is a technique that combines the general knowledge of LLMs with specific knowledge from a private dataset.

Applications of LLMs in eCommerce

Solving eCommerce's problems with LLMs

Content

- Data normalization
- Image tagging
- SEO
- Translation
- Content generation

UX

- Semantic search
- Conversational discovery
- Customer support

Solving eCommerce's problems with LLMs

Content

- Data normalization
- Image tagging
- SEO
- Translation
- Content generation

UX

- Semantic search
- Conversational discovery
- Customer support

Data Normalization

- Product metadata comes in all formats and completeness. Here the LLM takes the available metadata, and uses its **summarization** and **Q&A** capabilities to extract consistent metadata properties.

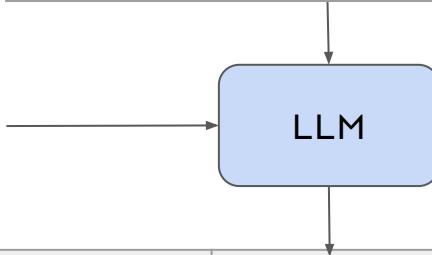
Data Normalization

Body HTML	Tags
<p><p>Inspired by the industrial style of the classic barn light, the Factory Light 6 Cord LED Pendant is pared down to sleek curves in handspun heavy gauge steel. A Schoolhouse take on the utilitarian, our meticulous attention to detail is apparent in the clean, deliberate lines and carefully hand-painted finishes.</p></p> <p><p>Suspended in space from a sleek vinyl cord, this pendant works in any space, whether its used to illuminate the whole room or a countertop. Each shade is handspun from steel and hand-painted in the Schoolhouse Factory.</p></p> <p><p>Assembled in our Portland, Ore. factory using ethically-sourced domestic and global components. A Schoolhouse Original.</p></p>	artisan-picks,best-of-both,blog-home-tour-sara-p arsons,damp_porch_patio,ETA18,ETA_18,F17,f actory,factory_pendants_chandeliers,findify-new, fuelmade,functional-artistry,Integrated LED Fixture,L-Nav18,LED_pendants_chandeliers,leng th_shade,Lighting,lighting-hardware,No_Return, Pendant,Pendant Fixtures,pendants_chandeliers,PFA23,PFA23-Li ghting,promo-010322-22,promo-051923,promo-0 52622-15,promo-052622-outdoorlighting,promo- 080222-20,Schoolhouse Signature,schoolhouse-wanders,Sept20LightPro mo,stories/the-beauty-of-schoolhouse-led-lightin g,traditional-industrial,Utilitarian

Data Normalization

Body HTML	Tags
<p><p>Inspired by the industrial style of the classic barn light, the Factory Light 6 Cord LED Pendant is pared down to sleek curves in handspun heavy gauge steel. A Schoolhouse take on the utilitarian, our meticulous attention to detail is apparent in the clean, deliberate lines and carefully hand-painted finishes.</p></p> <p><p>Suspended in space from a sleek vinyl cord, this pendant works in any space, whether its used to illuminate the whole room or a countertop. Each shade is handspun from steel and hand-painted in the Schoolhouse Factory.</p></p> <p><p>Assembled in our Portland, Ore. factory using ethically-sourced domestic and global components. A Schoolhouse Original.</p></p>	artisan-picks,best-of-both,blog-home-tour-sara-p arsons,damp_porch_patio,ETA18,ETA_18,F17,f actory,factory_pendants_chandeliers,findify-new, fuelmade,functional-artistry,Integrated LED Fixture,L-Nav18,LED_pendants_chandeliers,leng th_shade,Lighting,lighting-hardware,No_Return, Pendant,Pendant Fixtures,pendants_chandeliers,PFA23,PFA23-Li ghting,promo-010322-22,promo-051923,promo-0 52622-15,promo-052622-outdoorlighting,promo- 080222-20,Schoolhouse Signature,schoolhouse-wanders,Sept20LightPro mo,stories/the-beauty-of-schoolhouse-led-lightin g,traditional-industrial,Utilitarian

"How would you describe the {COLUMN} of this product?"



Type	Lighting	Style	Cord	Assembled
Pendant	Integrated, LED	Industrial, Barn, Sleek	Yes, Vinyl	Portland

Solving eCommerce's problems with LLMs

Content

- Data normalization
- **Image tagging**
- SEO
- Translation
- Content generation

UX

- Semantic search
- Conversational discovery
- Customer support

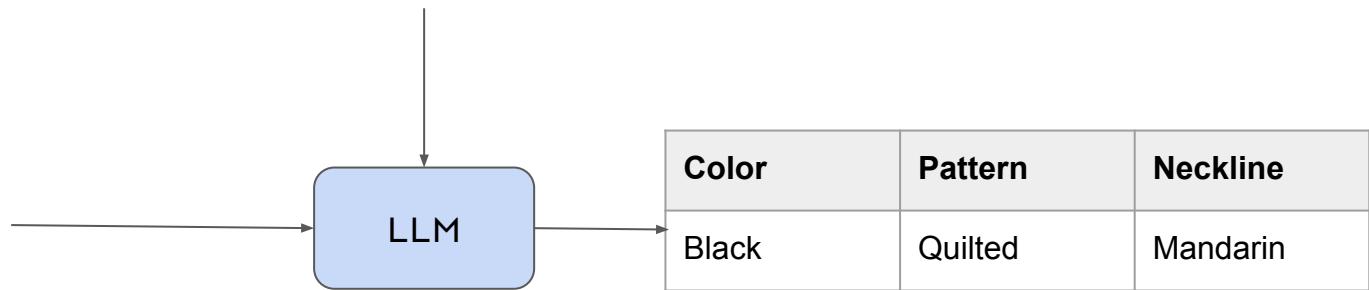
Image Tagging

- In many cases, rich product metadata is unavailable, however images are usually available.
- An LLM's multimodal input capability can be used to **describe** an image using specific tags in the prompt.

Image Tagging



"How would you describe the {COLUMN} of this product?"



Solving eCommerce's problems with LLMs

Content

- Data normalization
- Image tagging
- **SEO**
- Translation
- Content generation

UX

- Semantic search
- Conversational discovery
- Customer support

Search Engine Optimization (SEO)

- SEO can help products be discovered by search engines.
- We can again leverage LLM's **summarization** and **generative descriptiveness** capabilities to help us find tags that will likely match with the terms shoppers use to search for specific products.
- These tags can be used to enrich an on-site search engine, or used in HTML meta tags to aid in external search engine optimization.

Search Engine Optimization (SEO)

"What are common search terms that could make the banana more discoverable in search engines?"

Produce / Fruit /



Banana (each)

* Enter your address to view full delivery information.

[Sign In](#) [Set Address](#)

\$0.40 ea

[-](#) 1 [+](#)

[Add to cart - \\$0.40](#)

Search terms

Banana

Fresh fruit

Healthy Snack

Smoothie

Baking

Breakfast

Solving eCommerce's problems with LLMs

Content

- Data normalization
- Image tagging
- SEO
- **Translation**
- Content generation

UX

- Semantic search
- Conversational discovery
- Customer support

Translation

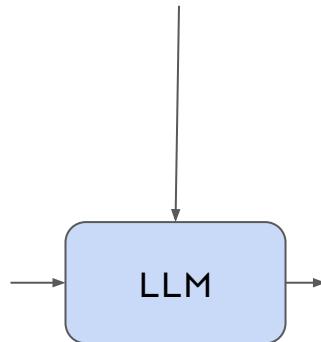
- With the growth of global eCommerce, stores have more to gain from selling to customers from other countries.
- Many LLM's like ChatGPT are **multilingual**, and can be used to easily translate entire product catalogs from one language to another.

Translation



グッドスマイルカンパニーから、『バンブレイブ バンブ
レイバーン』のDXブレイバーンの素晴らしいアクション
フィギュアが登場！アニメのアクション満載の瞬間を
捉るために完全可動式で、パーツの交換なしにブレ
イバーン形態からブレイブサンダー形態への変形が
可能です！4種類の顔パーツが含まれており、バーン
ブレードとバーンアックスを装備しています。バーンブ
レードのブレイズアップエフェクトパーツも付属してお
り、交換可能な手もあります。豊かな色彩を実現する
ために、透明パーツも使用されています。今すぐ注文
してください！

"Translate this product
description into English"



Good Smile Company brings us an awesome action figure of DX Bravern from "Bang Bang Bravern"! He's fully posable to capture action-packed moments from the anime, and he can transform from his Bravern form to the Brave Thunder form without parts replacement! Four face plates are included, and he's armed with the Burn Blade and Burn Axe; a Burn Blade Blaze Up effect part is included, as are interchangeable hands. Translucent parts are used for a richly colorful appearance, too. Order yours today!

Solving eCommerce's problems with LLMs

Content

- Data normalization
- Image tagging
- SEO
- Translation
- **Content generation**

UX

- Semantic search
- Conversational discovery
- Customer support

Content generation

- Building upon the previous two applications, we can leverage the **multimodal** and **multilingual generative** capabilities of LLMs to generate new or enhance existing product descriptions and even images.
- Prompt changes can allow for variations in voice, style, and expressiveness to match the store's brand and marketing goals.
- Social commerce marketing calls for a large volume of original content to post, which can be automated by LLMs.

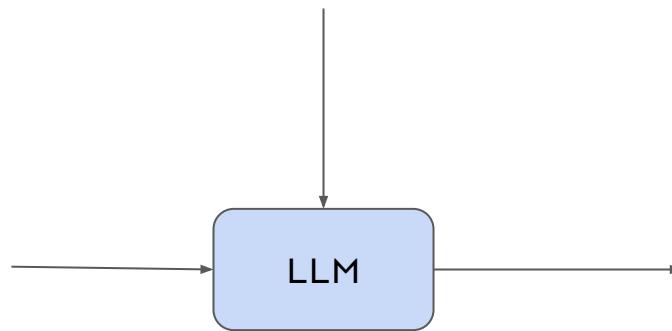
Content generation



Name: Two piece chaise patio love seat set

Description: This outdoor patio set is perfect for all seasons and waterproof for rainy weather. It's perfect for hosting outdoor bbqs, reading a nice book on a sunny day, or watching the kids play in the back yard. It's very durable, affordable, and comes with fast shipping. Dimensions: 74" W x 26.38" H x 48.43"

"Rewrite the product description to optimize for SEO, and also convey the values of quality, style, and value"



Planning a party, or just want some extra seating for the summer? This is the perfect outdoor chaise set. The durable wicker weave construction of the seating area has been designed for all seasons, giving you a comfortable spot for entertaining guests during spring and summer months, then a cozy place to relax in the falling temperatures during fall and winter. Enjoy BBQs, a good book, or watching your kids play from your new favorite spot.

Content generation

- LLMs can also be used to edit and generate images. By incorporating real product imagery with AI generated backgrounds and settings, LLMs can enhance the shopping experience by providing more engaging imagery.
- The imagery can also be tailored for different seasons, trends, and even personalized to the preferences of individual shoppers.



Content generation

- In the new mobile and social eCommerce world, content is what gets people engaged with your brand, and ultimately brings shoppers into your store.
- LLMs can be used to generate everything from blog posts to short videos, with interesting topics (best of lists, recipes, reviews, how-to guides) that can get potential shoppers to engage and lead to conversion.



Source: [One Ingredient Chef](#)

Solving eCommerce's problems with LLMs

Content

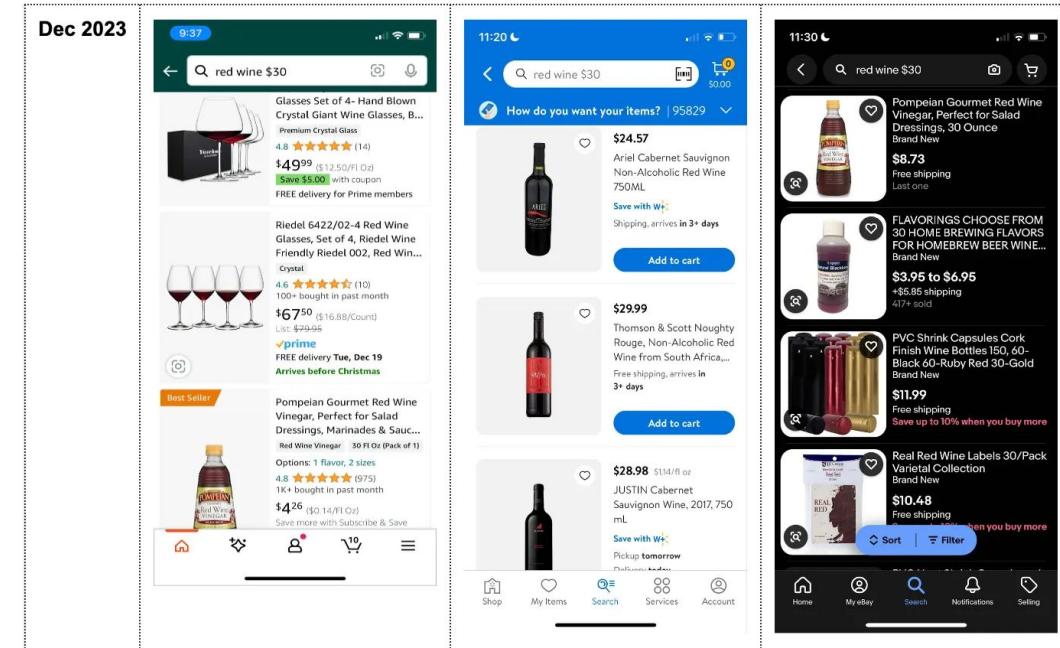
- Data normalization
- Image tagging
- SEO
- Translation
- Content generation

UX

- **Semantic search**
- Conversational discovery
- Customer support

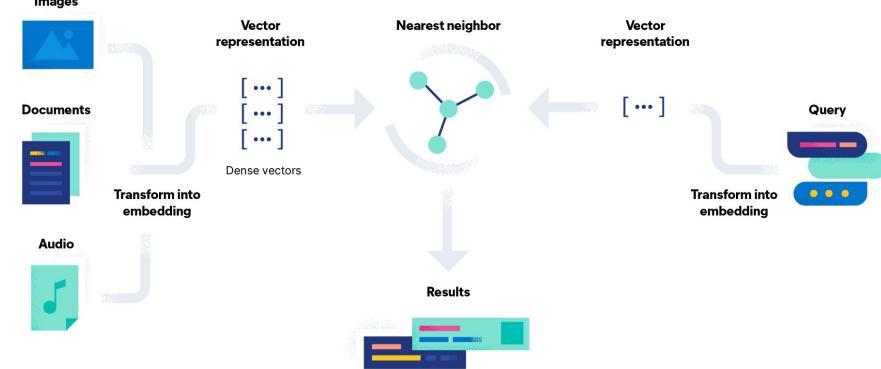
Semantic Search

- By using LLM's and embeddings of natural language queries, we can search for products not just by keywords, but by the semantic relatedness of products to what's expressed by the query.



Semantic Search

- A number of Vector DB platforms have emerged that offer search capabilities for vector embeddings.
- Many keyword-based search engines like Elasticsearch, Vespa, and Algolia now offer vector and hybrid search.
- Embeddings can come from a number of open-source and commercial models, like OpenAI, Huggingface, and so on.



Semantic Search

BODEN NEW DRESSES WOMEN GIRLS BOYS BABY & TODDLER **SALE**   

Home / Search results for "outfit for job interview"

Search results for "outfit for job interview"

510 items Department Category Type Sleeve Length Color Pattern Price Sort by: Relevancy



The image shows a grid of four outfit photos from Boden's website. From left to right: 1. A woman in a white blazer over a green top and striped pants, with a small circular icon containing a person icon and a plus sign below it. 2. A woman in a black blazer over a green top and black pants, with a plus sign below it. 3. A woman in a purple and yellow patterned dress, with a plus sign below it. 4. A woman in a blue and white patterned dress, with a green circular icon containing a speech bubble icon and a plus sign below it.

Solving eCommerce's problems with LLMs

Content

- Data normalization
- Image tagging
- SEO
- Translation
- Content generation

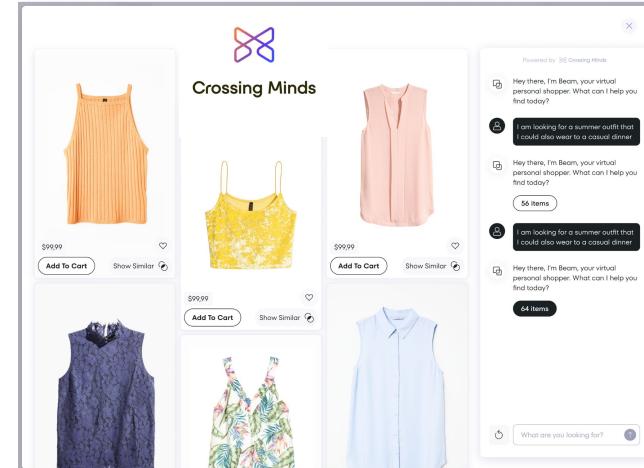
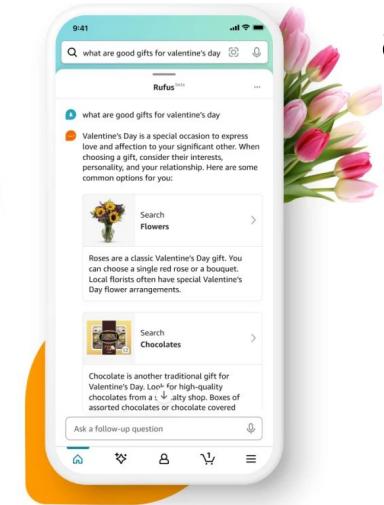
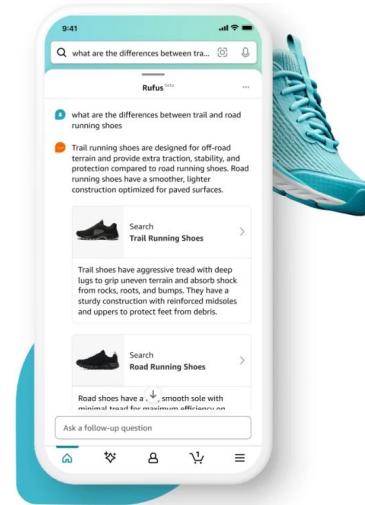
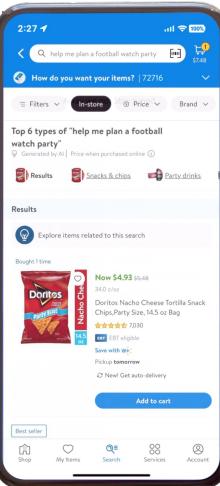
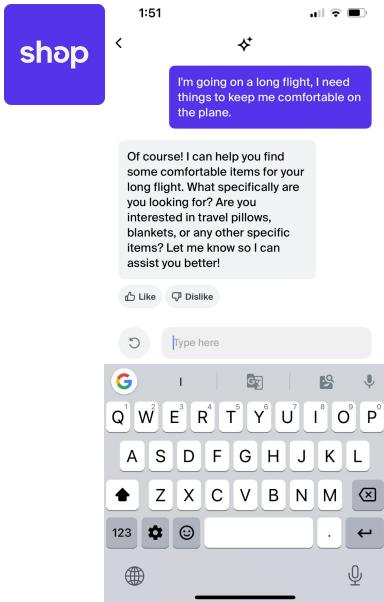
UX

- Semantic search
- **Conversational discovery**
- Customer support

Conversational discovery

- Since ChatGPT, conversational “chatbot” interfaces have been hyped as the next big thing in eCommerce discovery. Many major retailers have launched a conversational shopping assistant feature, but we are still in the early stages.
- LLMs offer not just semantic search, but also multi-turn, interactive product discovery, with a UX that is familiar to mobile text messaging app users.
- Challenges:
 - Balancing the accuracy of search with the inspiration of generative AI.
 - Safely handling non product-related queries (customer support, discounts, spam/abuse).

Conversational discovery



Solving eCommerce's problems with LLMs

Content

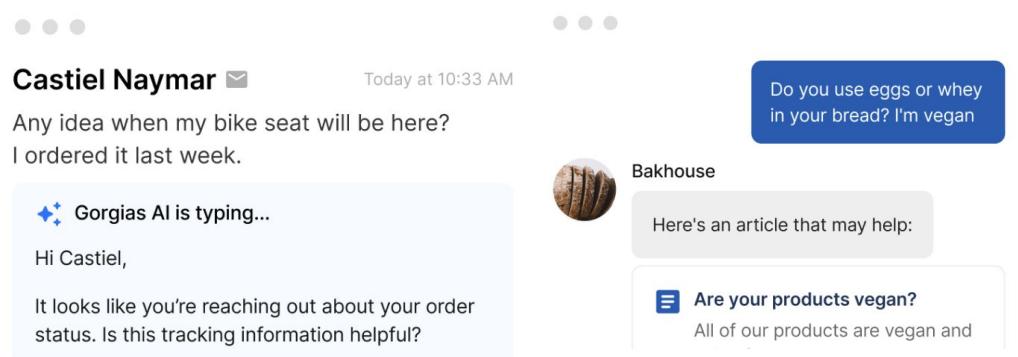
- Data normalization
- Image tagging
- SEO
- Translation
- Content generation

UX

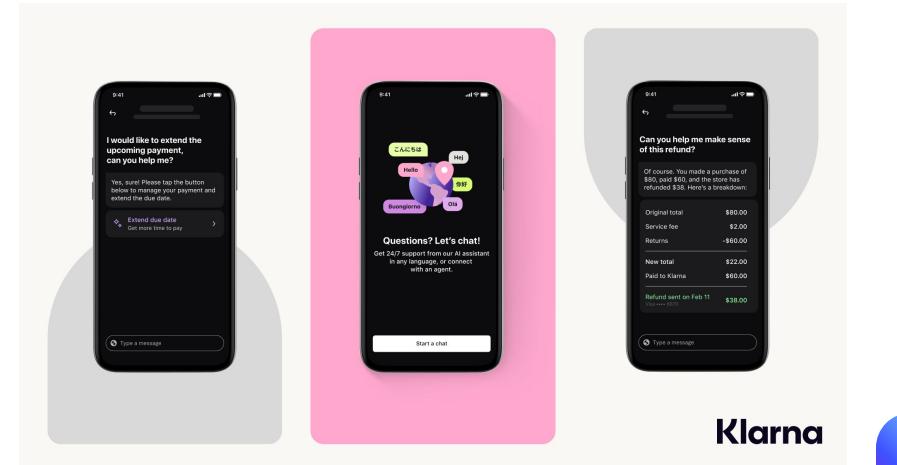
- Semantic search
- Conversational discovery
- **Customer support**

Customer support

- Customer support is a costly, labor-intensive feature that is hard to offer 24/7.
- LLM-based Customer Support can reliably offer assistance for common requests around the clock, giving a better customer experience.
- Klarna recently [announced](#) that their AI Assistant is handling two-thirds of their customer service chats, resolving issues 5x faster, and will drive \$40M in profits in 2024.



 gorgias



Solving eCommerce's problems with LLMs

Content

- Data normalization
- Image tagging
- SEO
- Translation
- Content generation

UX

- Semantic search
- Conversational discovery
- Customer support

Thank You!