Google DeepMind

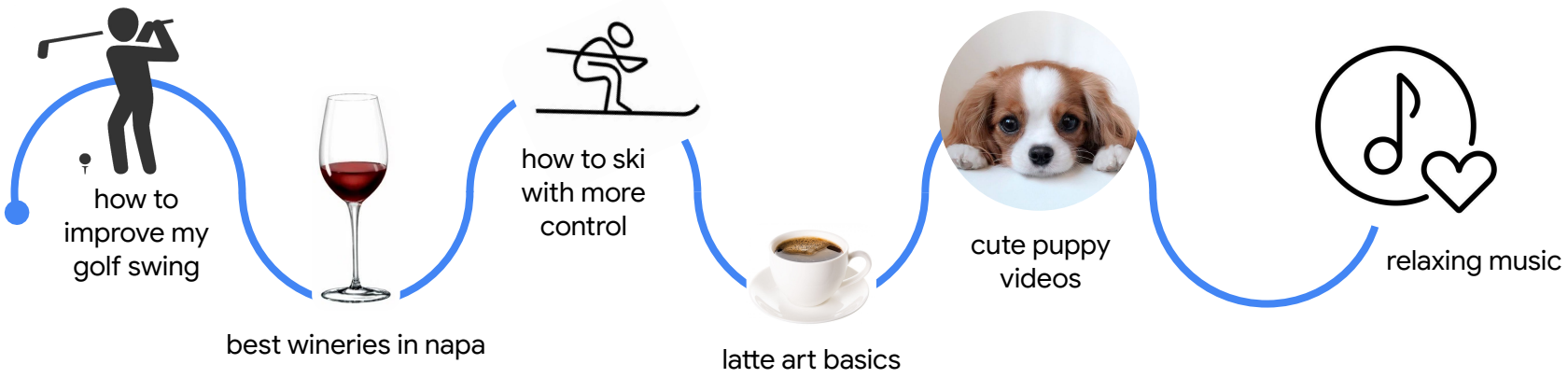
ALLY: LLM-based Assistive Recommendation

Konstantina Christakopoulou
Senior Research Engineer, Tech Lead


How can I align my work on what matters to me?
The real-life things people pursue

Vision

Build a **companion recommender** (ally) **a**ssisting the user‘s **l**ife**l**ong journe**y.**



Current RecSys are more focused on the short-term

Entertainment/Hobbies (50%)
Gaming, photography, music, film

Learning new skills (34%)
Language, coding, certificates, make music/videos/games

Travel
Planning, inspirations, camping, things to do

Wellbeing
Mental health, relationships, healing, self-improvement

DIY projects
Sewing, knitting, quilting, woodworking, interior

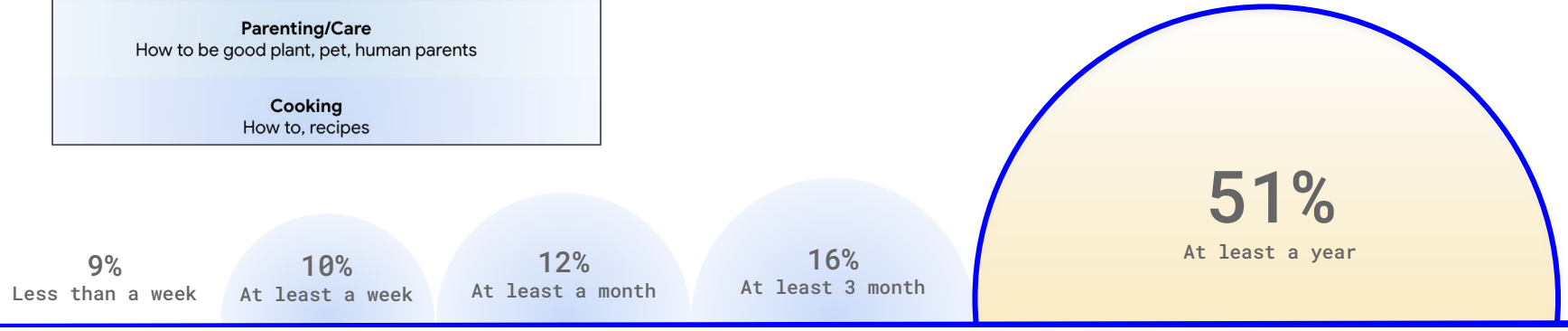
Maintenance/Repair
Homes, cars, hardwares

Fitness
Exercise, yoga, meditation

Parenting/Care
How to be good plant, pet, human parents

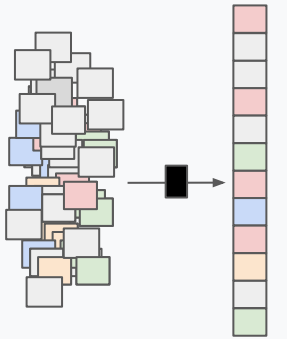
Cooking
How to, recipes

*But, we know that users pursue their interests for a **long time**.*



And, current RecSys don't offer interpretability and control

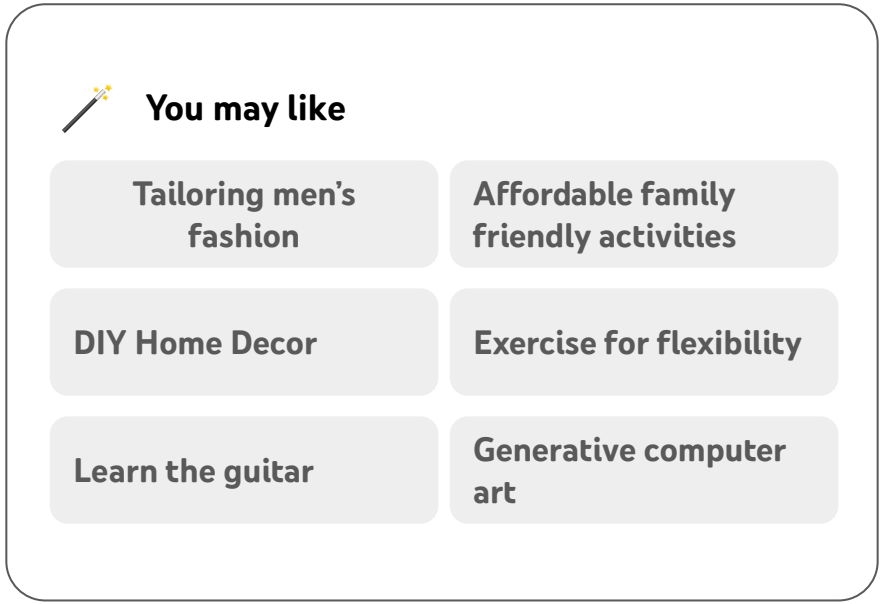
Recommendations today **mixes user interests** in a single feed.



User Profile

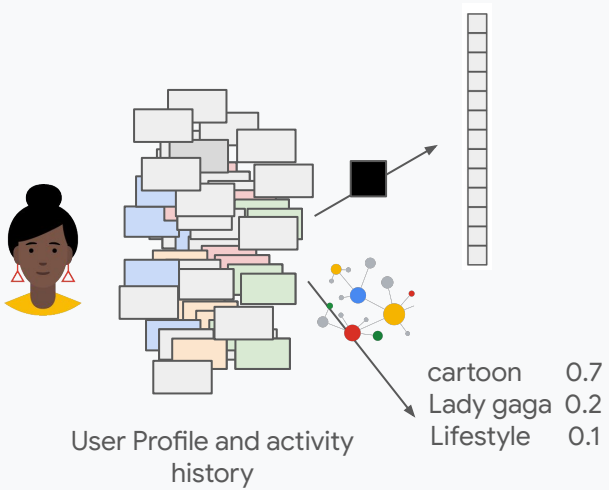
Mixed Recs

*How might we go about **steerable recommendations** on interest journeys?*



Also, current RecSys don't speak the same language as users

User interests today are captured as **embeddings** or too broad **knowledge graph entities**

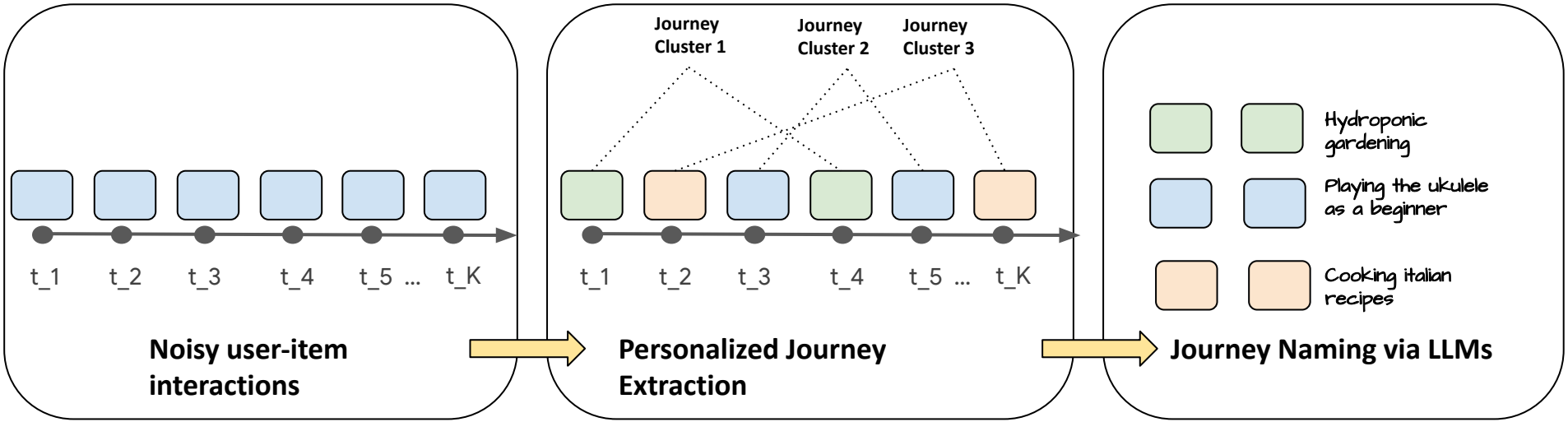


*But, users use **nuanced language** to describe their journeys in a **much more** personalized way*

How to move towards **assistive recommenders** that

- Account for long-term journeys, evolving in a personalized fashion
- Speak the same language as users
- Offer steerable journey recommendations

In a nutshell

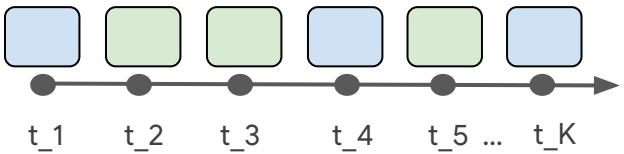


Interest journey service for user interest profile

LLMs for User Interest Journeys. K Christakopoulou et al. arXiv

Global Interest Extraction

Noisy user-item interactions



Annotate each item using **pre-defined** interest clusters. E.g.,

- Co-occurrence based,
- Semantic similarity based

The granularity of the pre-defined interest cluster determines how nuanced the user interest can be.

But, the same item can be a part of different user journeys

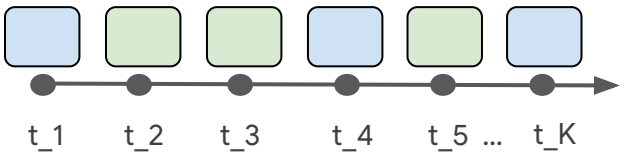
Hydroponic gardening

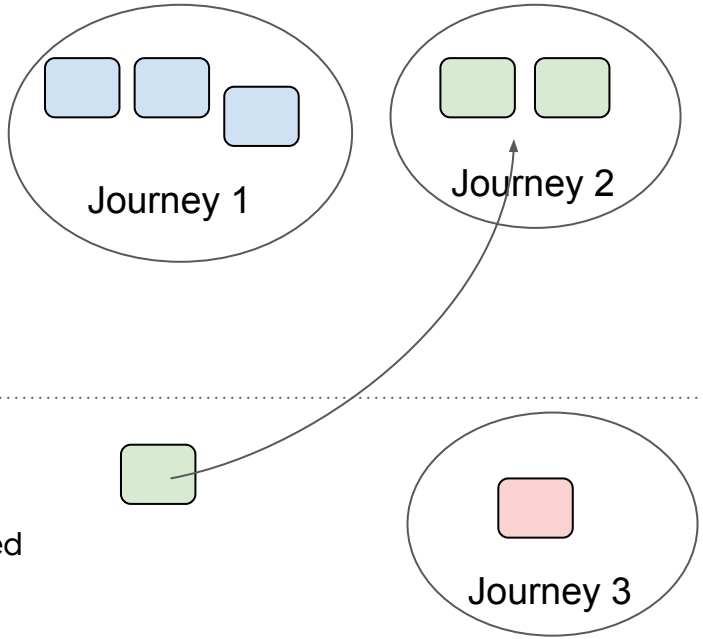
Italian pizza making

Growing herb and vegetables in my patio

Personalized Interest Extraction

Noisy user-item interactions





Personalized clustering update:
When an item's representation has **high cosine similarity** with the **journey cluster**, it should be added to this cluster. Otherwise, a new cluster is created.

Personalized vs Global Interest Extraction

Baselines: Multi-modal similarity topic clusters, Co-occurrence clusters

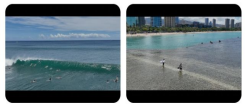
MULTIMODAL CLUSTERS

6 documents

Surfing Big Bowls 🔥 (June 27, 2022) 4K | BARRELS AND WIPEOUTS AT ALA MOANA BOWLS! May 3rd, 2022 | South Shore Surf (June 30, 2022) 4K | Surfing Canoes (June 23, 2022) Boys & Girls 4K | Surfing Threes (June 21, 2022) 4K | Surfing Publics Raw POV (June 27, 2022) 4K

CO-OCCURRENCE CLUSTERS

2 documents



Surfing Big Bowls 🔥 (June 27, 2022) 4K | South Shore Surf (June 30, 2022) 4K

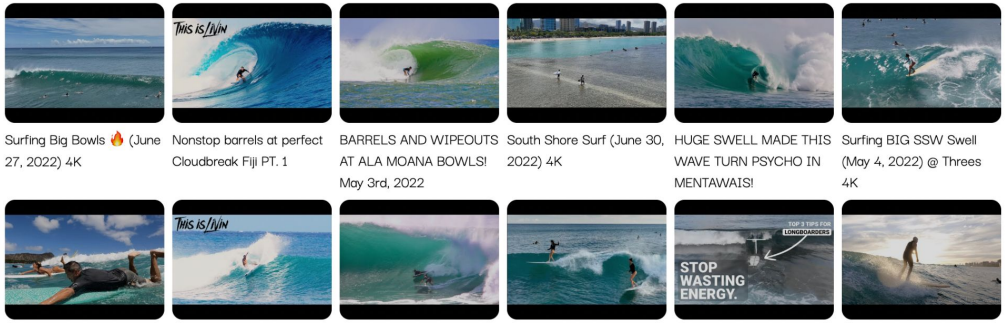
2 documents

Nonstop barrels at perfect Cloudbreak Fiji PT. 1 | As good as it gets for Waikiki (Bowls) Attacked by crackheads!

PERSONALIZED JOURNEY EXTRACTION

SURFING, SURF, ALA MOANA, WAIKIKI, THIS IS, MOANA BOWLS, HONOLUA BAY, THREES, MENTAWAI, IS LIVIN, BOWLS, LIVIN, MOANA, ALA, FIJI, THIS, CLOUDBREAK, LONGBOARD, IS, HOW, SURF SPOT, 2022, SOUTH SHORE, HAWAII, SURFING THREES, PUBLICS, PUBLICS SURF, HAPPY LIFE, SWELL, FIGHTS, HAWAII HAPPY, SPOT, OAHU, MENTAWAI ISLANDS, KOA ROTHMAN, HAPPY, LIFE, EIMYS, HONOLUA, BAY, NATHAN FLORENCE, TURTLE ROLL, LONGBOARD SURFING, CLOUDBREAK FIJI,...

12 documents



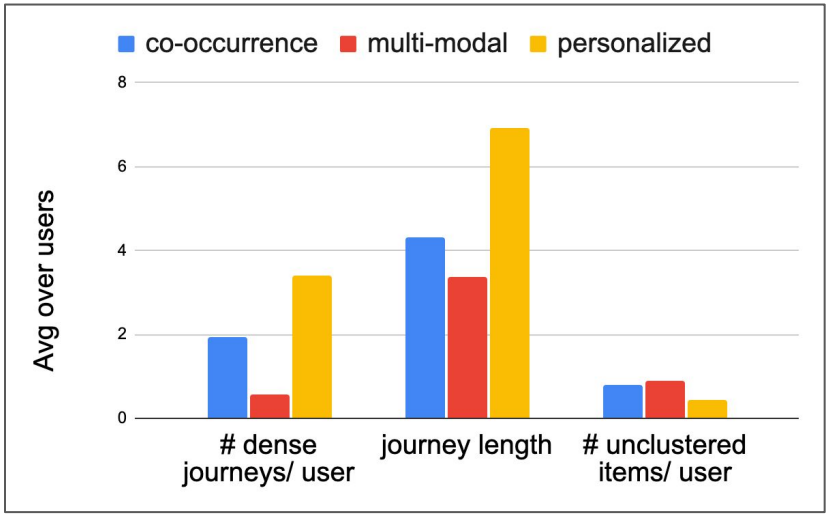
Surfing Big Bowls 🔥 (June 27, 2022) 4K | Nonstop barrels at perfect Cloudbreak Fiji PT. 1 | BARRELS AND WIPEOUTS AT ALA MOANA BOWLS! May 3rd, 2022 | South Shore Surf (June 30, 2022) 4K | HUGE SWELL MADE THIS WAVE TURN PSYCHO IN MENTAWAIS! | Surfing BIG SSW Swell (May 4, 2022) @ Threes 4K

Surfing Canoes (June 23, 2022) Boys & Girls 4K | As good as it gets for Waikiki (Bowls) Attacked by crackheads! | Barrels and Brawls - Drop-in at Honolua Bay leads to collision (January 1, 2019) | Surfing Threes (June 21, 2022) 4K | How to Pass the Break on a Longboard | How to Surf | Surfing Publics Raw POV (June 27, 2022) 4K

Personalized yields the **longest**, **most coherent** journeys

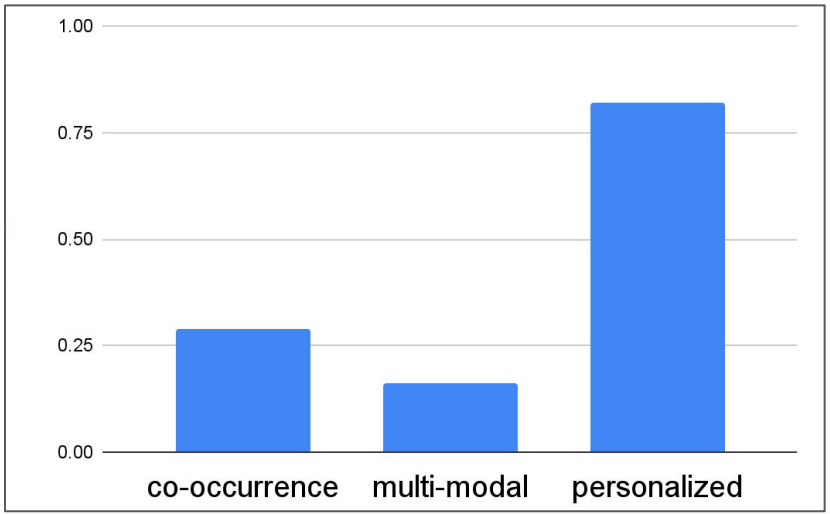
But, how do we even know the extracted journeys are good?

Insight: Granularity affects recall
→ Proxy granularity metrics

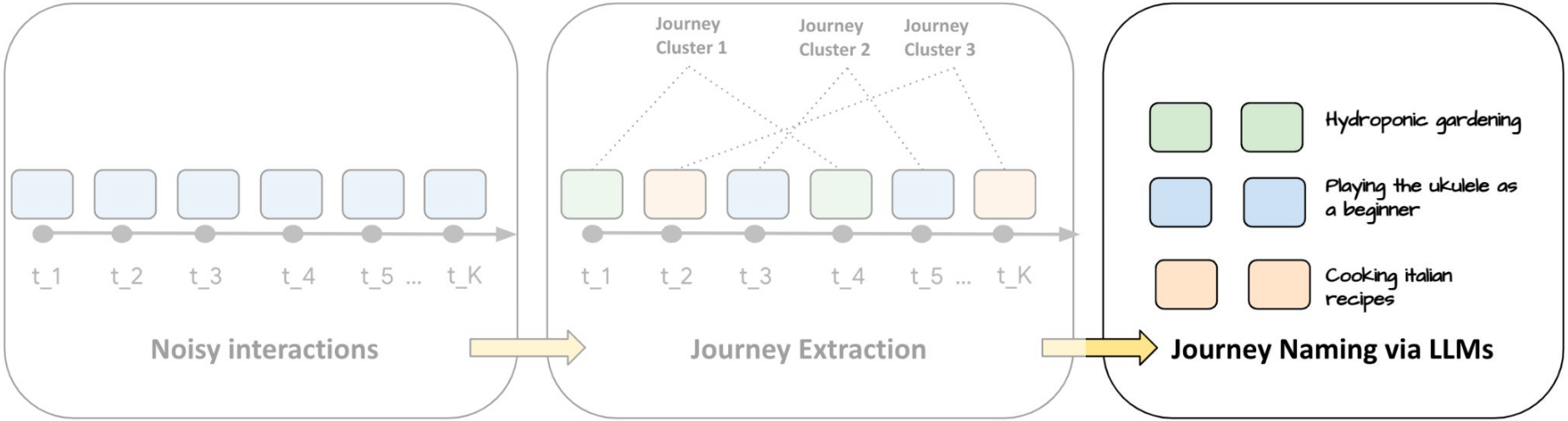


Personalized yields the **most dense,** and **longest** journeys

Idea: Rely on **expertly curated playlists with high episodicity** → Ask the algorithm to divide into journeys.



Personalized leads to **182.7% increase in recall**!


t_1
t_2
t_3
t_4
t_5 ...
t_K
Noisy interactions
Journey Cluster 1
Journey Cluster 2
Journey Cluster 3
t_1
t_2
t_3
t_4
t_5 ...
t_K
Journey Extraction
Hydroponic gardening
Playing the ukulele as a beginner
Cooking italian recipes
Journey Naming via LLMs

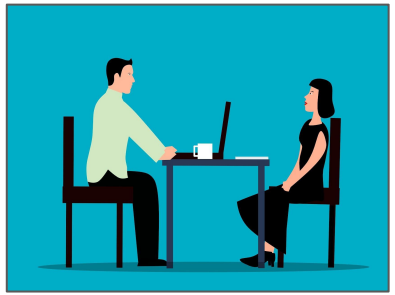
Of course, we could ask an LLM to provide the interest name

Prefixes	I interacted with items with titles:	So, I may be interested in
Examples (optional)	“David Graeber - Debt, service, and the origins of capitalism”, “Mark Fisher and Our Contemporary Moment: Is There Still No Alternative”, “The Third Industrial Revolution: A Radical New Sharing Economy”	Critical theory of society and contemporary culture
	"ruth bader ginsburg: legal pioneer", "makers.com; who is kamala harris? a look at her background and career in politics" "malala yousafzai on the power of education" "glamour; oprah winfrey opens up about journey to become first woman to own & produce her own talk show" "time; madeleine albright, first female secretary of state" "msnbc; the first woman in space - valentina tereshkova i the cold war"	Trailblazer women
	Ask model for a new journey	
Inference	"4 ways to use hardware synthesizers", " Analog Synths with Ableton Push | Set Up & Creative Ideas", "HOW TO BUILD A SYNTHESIZER SETUP with Circuit Tracks + Korg Minilogue & Volca NuBass "	Synth music production

- Can we do better?

- YES! Aligning with the right data

What data to use to align LLMs to user interest journeys?



User Interviews

Expert curated collections

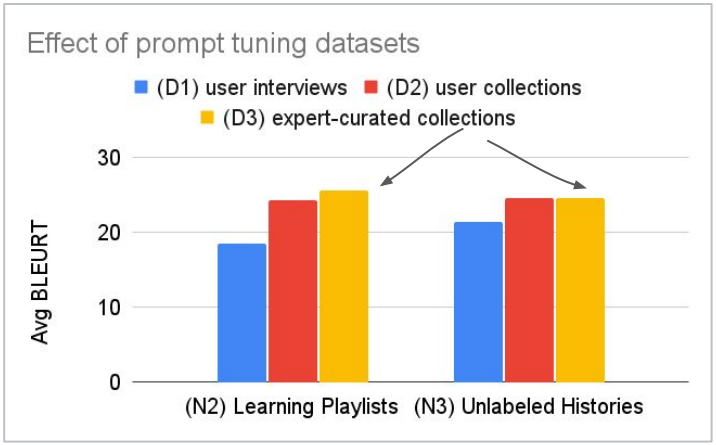
Saved personal collections

Online playlists

Small, higher-quality data ~O(100)
Ideal for **prompt tuning**

Larger scale, noisy data O(10K)
For **fine tuning**

What data to use to align LLMs to user interest journeys?



Ideal for prompt tuning

Expert curated collections perform the best

Online playlists

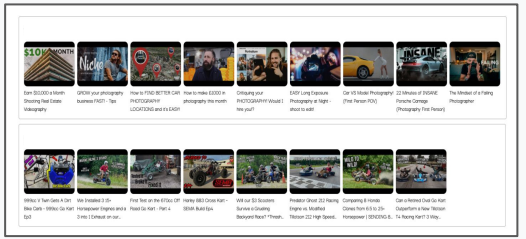
For fine tuning

How to evaluate effectiveness of journey naming?

(N1) Setup: 100 Expert curated collections

(N2) Setup: 10K learning high episodicy playlists

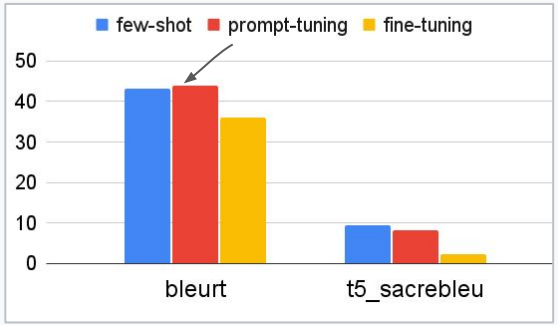
(N3) Setup: 2K user histories over a month ⟶ ~11K journeys



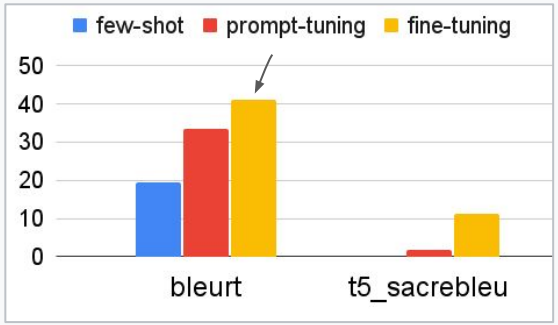
Three experimental evaluation setups

Which alignment technique performs the best?

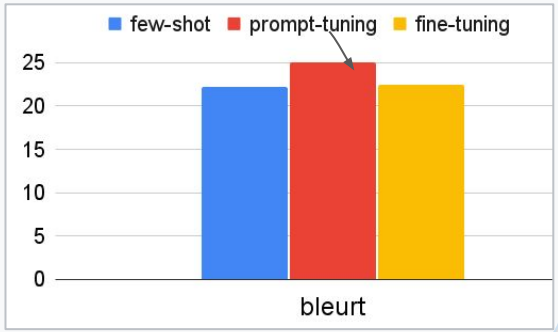
(N1) Setup: 100 Expert curated collections



(N2) Setup: 10K learning high episodicy playlists



(N3) Setup: 2K user histories over a month

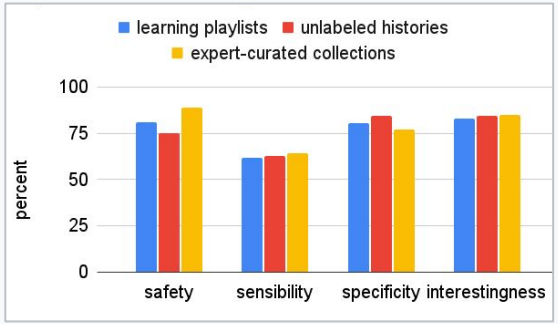


For in-domain evaluation, the technique using **the in-distribution data** is the highest performing one

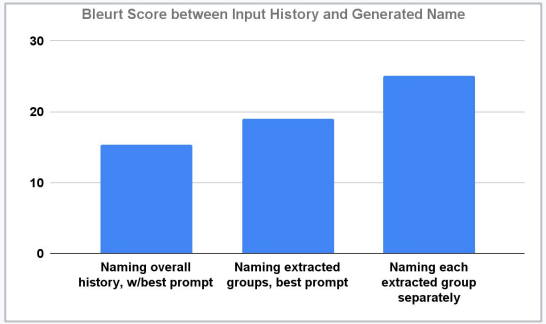
Prompt tuning has the **best generalization capability** out-of-domain

Other key insights

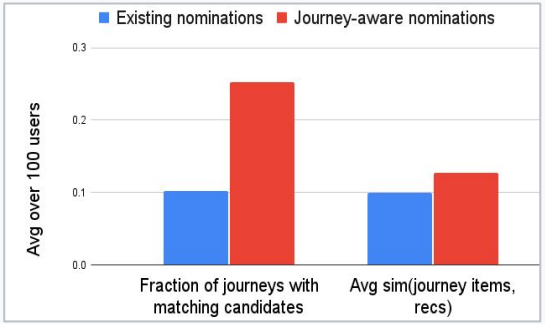
Inputs largely determine response safety



Journey extraction improves naming quality

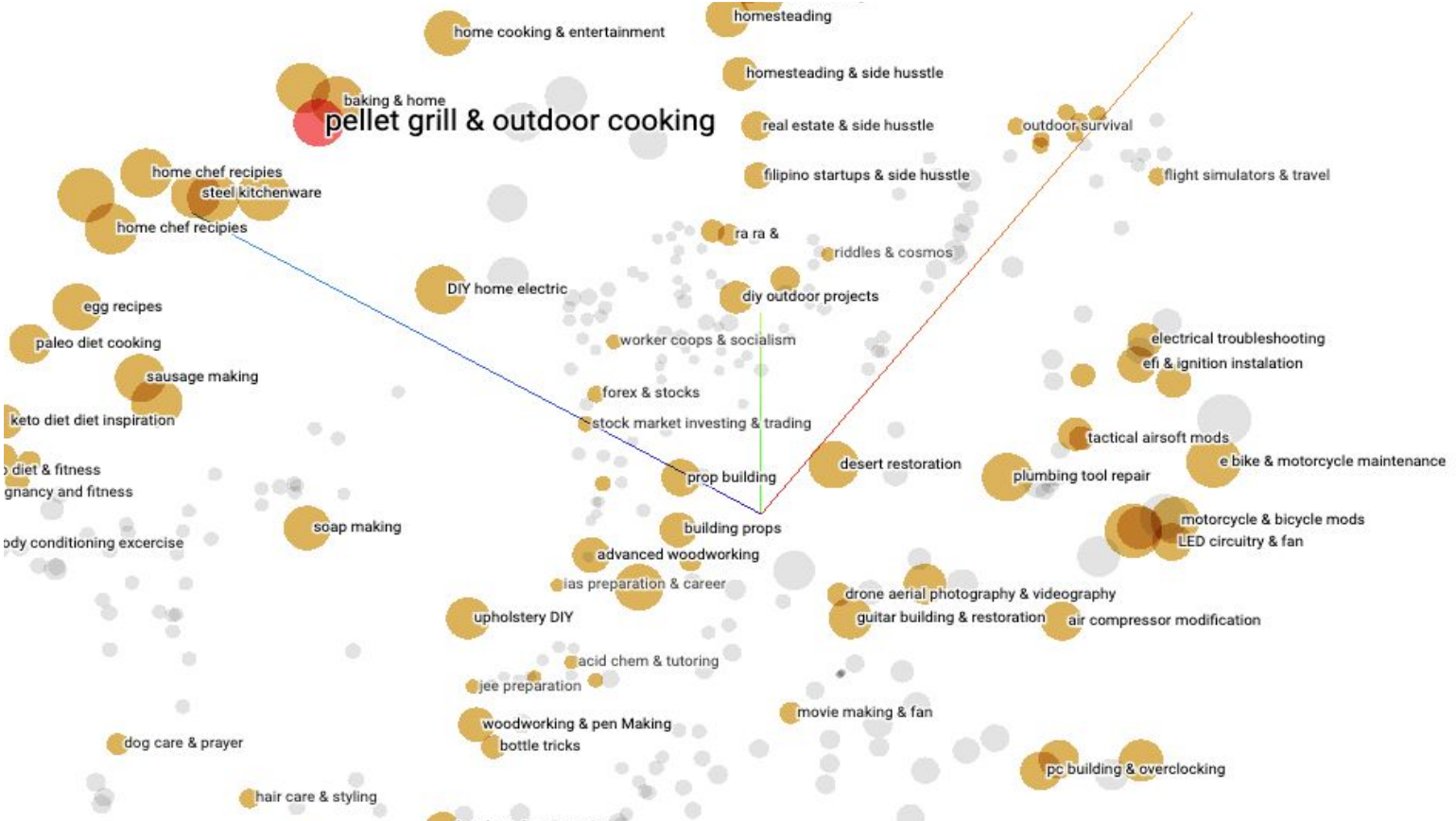


Journey-aware recs better support user interests



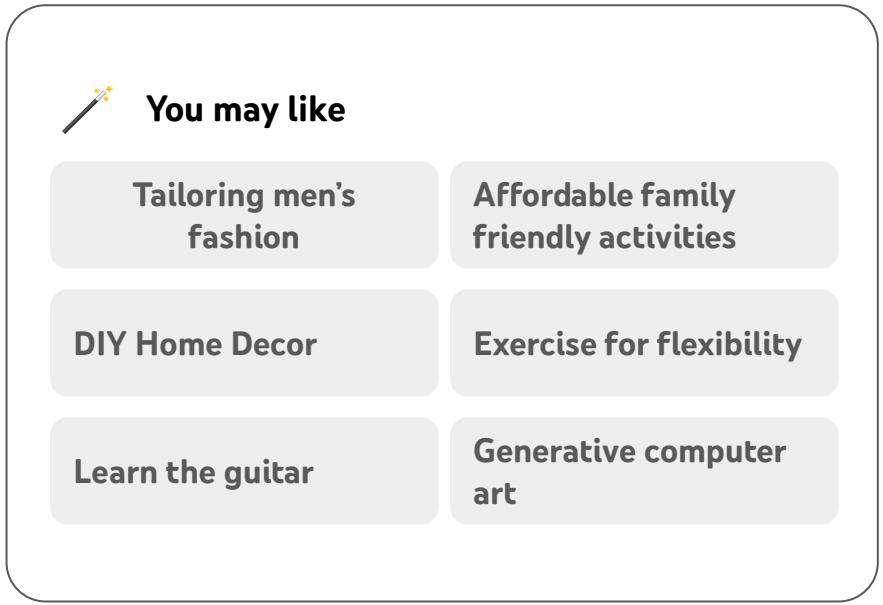
LLMs can unlock better user understanding

LLMs generates **nuanced and informative** description of real user journeys.



Leveraging journeys to provide steerable recommendations

Frontend experience allowing users to **delve deeper** into their interest journeys, or **explore related** journeys



Statistically significant A/B Experiment Results

Engagement	+0.04%
Abandonment	-0.05%
Engaged Users	+2.63%
Recs diversity	+0.15%


Are we done?
No!

Towards Companion Recommenders Assisting Users' Long-Term Journeys. K Christakopoulou and Minmin Chen. RecSys 2023

Companion Recommenders

How would a companion recommender, capable of accompanying users through their real-life interests, needs and goals, look like?

Create magical conversational journey-aware experiences?

Accompany the user to their evolving journey, assessing users' progress?

Plan at the abstract level of journeys and sub-journeys, while providing interpretability & control?

...