



# Multi-Stakeholder aware

# Recommender Systems

Bereket A. Yilma



# Session 3



- A Multi-Stakeholder aware RecSys
- Formulating the RecSys Problem? →(A framework)
- Modern RecSys Paradigms
  - ◆ Large Language Models as RecSys (Zero & Few shot)
- Common Issues and Challenges in RecSys



RECOMMENDER  
SYSTEMS



# Formulating a RecSys Problem



Task: Design a Personalised **Visual Art Recommendation** engine for NG/ Louvre

# Formulating a RecSys Problem



An illustration of a person wearing overalls and a white t-shirt, mopping a floor. The floor has a large blue circle with stars and several binary code patterns (0s and 1s) scattered across it. A bucket and cleaning supplies are nearby.



- Sort
  - Filter
  - Recommend



## **Understanding the context of the problem!**

# Formulating a RecSys Problem



Why does understanding the context matter?



Task: Design a Personalised **Visual Art Recommendation** engine for NG/ Louvre

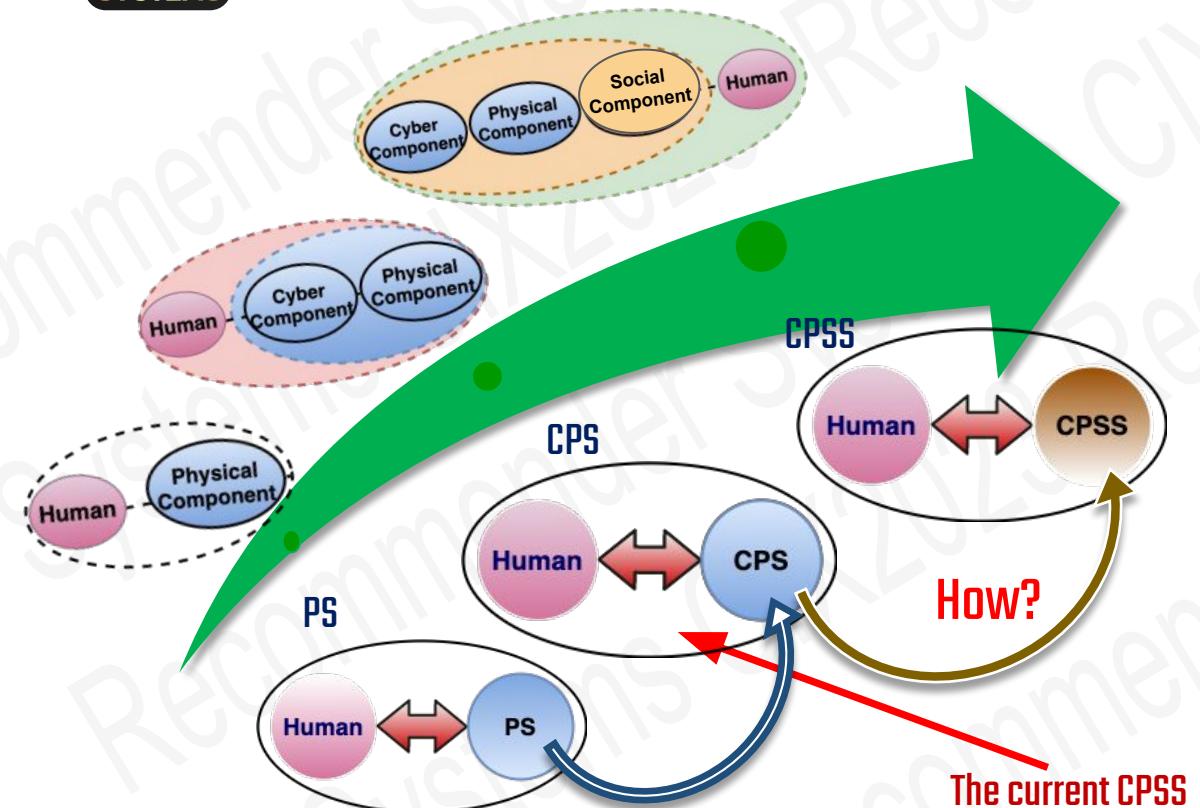
# Formulating a RecSys Problem



- Points of Interest
- Sensors, cameras, (CPS)
- Smart devices (CPSS)
- Visitors



# Cyber-Physical-Social System (CPSS)

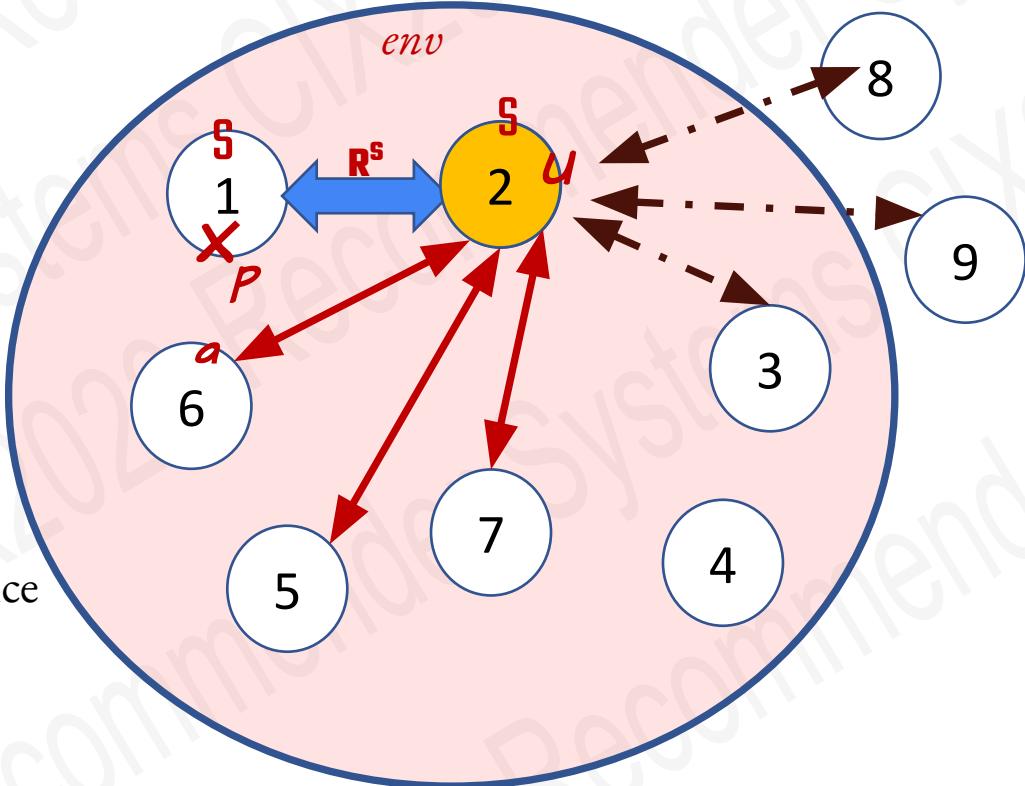


# Formulating a RecSys Problem



- Smart system environment  $env$ ,
- **Personalisation** is a function of a social component **S** of a system.
- **Personaliser**( $x_{pa}$ );
- **User**( $U$ )
- **Crowd**( $Cr$ ): direct influence
- **Context elements**( $Cx$ ): indirect influence

$$Pa^{coss} = f(u, x_{pa}, cr, cx, env)$$



# Formulating a RecSys Problem

Personalisation in exhibition areas for a user  $u$  can be formalised as a function of

- The user  $u$ ,  Visitor  $vs$
- The personaliser  $X_{pa}$ ,  Mobile guide  $mg$
- The crowd  $Cr$ ,  Crowd of other Visitors  $cr^{vis}$
- The context elements  $cx$
- The Smart environment  $env$   Exhibition area  $ex$

$$pa = f(u, X_{pa}, cr, cx, env) \longrightarrow pa^{Exhib} = g(vs, mg, cr^{vis}, ex)$$

# Formulating a RecSys Problem

$$pa^{Exhib} = g(\mathbf{vs}, \mathbf{mg}, \mathbf{cr}^{vis}, \mathbf{ex})$$

## ➤ User (Visitor):

- Preferences/interests
- Time (Limited availability)
- Crowd tolerance
- Visiting style
- Fatigue
- (Disability)
- (Age)



# Formulating a RecSys Problem



- **The Ant visitors:** spend a long time observing all exhibits moves close to the walls and the exhibits avoiding empty space.



- **The Fish visitors:** walk mostly through empty space making just a few stops and sees most of the exhibits but for a short time.



- **The Grasshopper visitors:** see only exhibits they are interested in and walk through empty space and stay for a long time only in front of selected exhibits.



- **The Butterfly visitors:** frequently change the direction of the tour route, usually avoiding empty space. They sees almost all exhibits, but times vary between exhibits.

# Formulating a RecSys Problem

$$pa^{Exhib} = g(\mathbf{vs}, \mathbf{mg}, \mathbf{cr}^{vis}, ex)$$

- **User (Visitor):**
    - Preferences/interests
    - Time (Limited availability)
    - Crowd tolerance
    - Visiting style
    - Fatigue
    - (Disability)
    - (Age)
  - **Exhibition area/Curator**
    - Preferences/interests
    - Time (Opening hours)
    - Crowd capacity
  - **Crowd**
    - Preferences/interests (Popularity of items)
    - Congestion (Crowd size)
- 
- 

Ant      Butterfly      Fish      Grasshopper

Najbri et al. 2014

# Multi-Stakeholder aware RecSys



Curator-visitor tradeoff

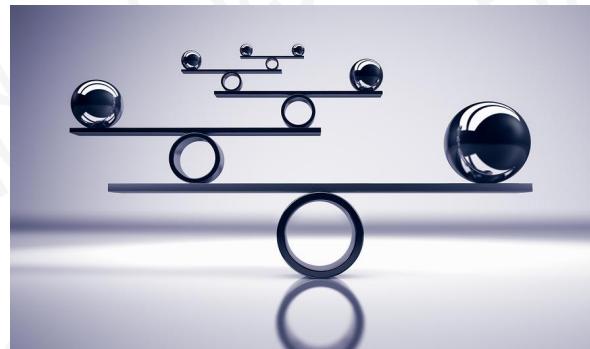


# Multi-Stakeholder aware RecSys



$$pa^{Exhib} = g(vs, mg, cr^{vis}, ex)$$

The personaliser needs to make the best possible compromise to satisfy the **Objectives** of the **co-existing stakeholders** while respecting environmental **Constraints**.



**Constrained multi-objective optimization problem**

# The RecSys pipeline: A case-study approach



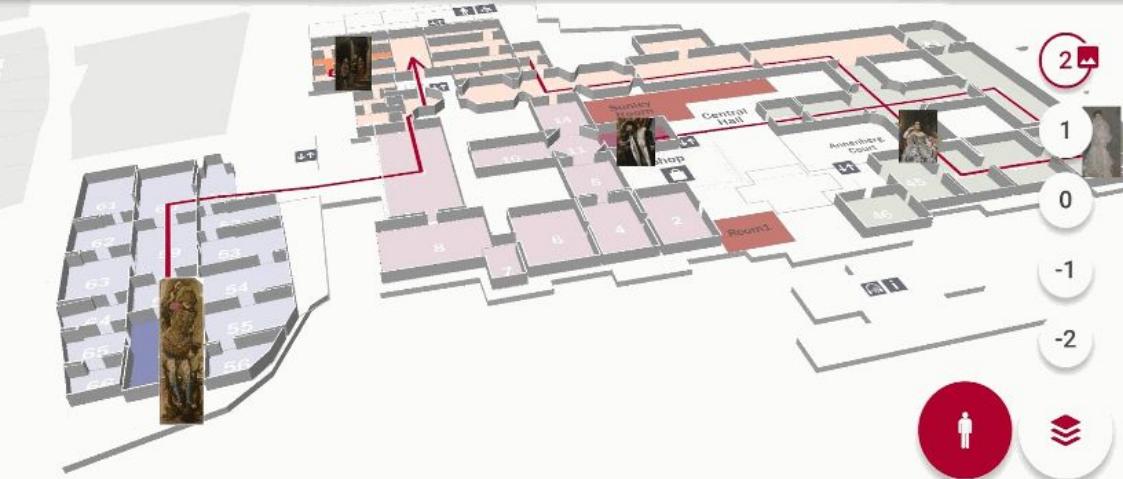
**Task:** Design a Personalised **Visual Art Recommendation** engine for the National Gallery, London

# Multi-Stakeholder aware RecSys

## 1. POI (painting) Recommendation

## 2. Path Recommendation

### Contemporary Style and Fashion





# The typical RecSys Pipeline

Data  
Pre-processing



Model  
Training

Post  
Processing

Evaluation



- Sort
- Filter
- Recommend



# Multi-Stakeholder aware RecSys



Data  
Pre-processing



New Visitor



## Query User (Profiling)

1. Rate few paintings
2. Popular paintings
3. Visiting style
4. Available time ...



# Multi-Stakeholder aware RecSys

Data  
Pre-processing



Task  
Personalised  
Recommendation

Model  
Training

Good representation of  
the data!

$$R^{m \times m}$$

1	0.77	0.57	0.54	0.37	0.46	0.45	0.44	0.46	0.37	0.63	0.66	0.59	0.54	0.59	0.52
0.77	1	0.69	0.68	0.54	0.56	0.61	0.53	0.5	0.46	0.74	0.85	0.66	0.58	0.67	0.6
0.57	0.69	1	0.92	0.34	0.39	0.45	0.4	0.43	0.45	0.48	0.59	0.57	0.62	0.67	0.59
0.54	0.68	0.92	1	0.37	0.38	0.5	0.44	0.42	0.41	0.47	0.59	0.55	0.6	0.66	0.58
0.37	0.54	0.34	0.37	1	0.62	0.55	0.51	0.48	0.52	0.55	0.62	0.42	0.38	0.39	0.39
0.46	0.56	0.39	0.38	0.62	1	0.58	0.54	0.49	0.6	0.53	0.7	0.47	0.38	0.39	0.34
0.45	0.61	0.45	0.5	0.55	0.58	1	0.7	0.39	0.45	0.57	0.65	0.58	0.47	0.54	0.45
0.44	0.53	0.4	0.44	0.51	0.54	0.7	1	0.28	0.39	0.44	0.57	0.6	0.48	0.47	0.39
0.46	0.5	0.43	0.42	0.48	0.49	0.39	0.28	1	0.65	0.49	0.52	0.47	0.37	0.4	0.37
0.37	0.46	0.45	0.41	0.52	0.6	0.45	0.39	0.65	1	0.5	0.57	0.47	0.41	0.44	0.42
0.63	0.74	0.48	0.47	0.55	0.53	0.57	0.44	0.49	0.5	1	0.82	0.53	0.43	0.58	0.51
0.66	0.85	0.59	0.59	0.62	0.7	0.65	0.57	0.52	0.57	0.82	1	0.61	0.53	0.66	0.61
0.59	0.66	0.57	0.55	0.42	0.47	0.58	0.6	0.47	0.47	0.53	0.61	1	0.7	0.53	0.45
0.54	0.58	0.62	0.6	0.38	0.38	0.47	0.48	0.37	0.41	0.43	0.53	0.7	1	0.53	0.46

- If a user likes painting A find paintings B, C, D that are similar to A.

## 1. Profiling

New Visitor



- List of paintings rated  $P^u = \{P_1, P_2, \dots, P_n\}; P^u \in P$
- Rating of  $P^u$ ,  $W^u = \{w_1, w_2, \dots, w_n\}$
- Available time  $T_{ava}$
- Visiting style (Ant, butterfly, fish, grasshopper)
- Crowd tolerance  $C_t(u)$
- $(\beta, \lambda, \varepsilon)$  Popularity, Fatigue and Diversity tolerance

Museum



- Similarity matrix from LDA/BERT/ResNet
- Opening hours
- Crowd capacity
- Curated Stories
- Rules regarding movement in the physical space.

## 1. POI recommendation → 1.1 Matching User Preferences



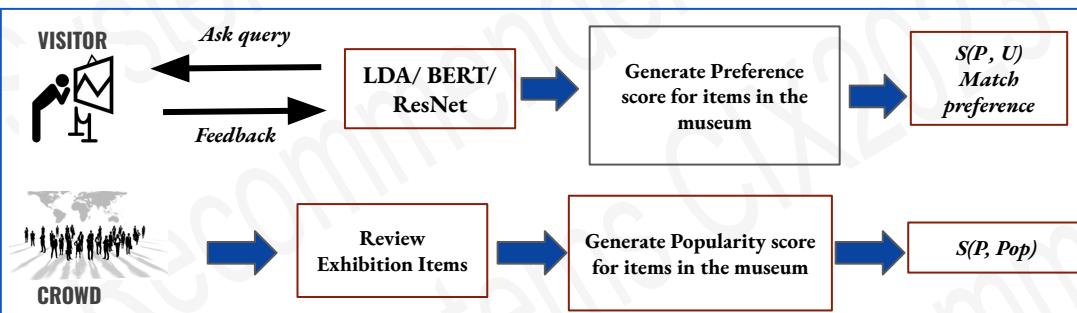
- In addition to unique personal preferences users also have different tendency to be interested in visiting famous paintings. Hence, we introduce a popularity score  $S(p, Pop)$  for all the paintings in the dataset. This score is based on public review (**Crowd**) from National Gallery website.

- By taking into account the ***preference of the user*** and also the ***crowd opinion*** we generate an aggregate preference score  $S(P)$  for the paintings in the dataset.

$$S(P) = \alpha S(P, U) + \beta S(P, Pop)$$

- $\beta$  is user provided hyper parameter determining user's interest to see popular items.

$$\alpha = 1 - \beta$$



## 1. POI recommendation →

### 1.2 Matching Curator's Goal



- The exhibition curator might have different goals related to the point of interests to be presented for visitors.
- In this case study the curator's goal is to increase the number curated stories presented to visitors.

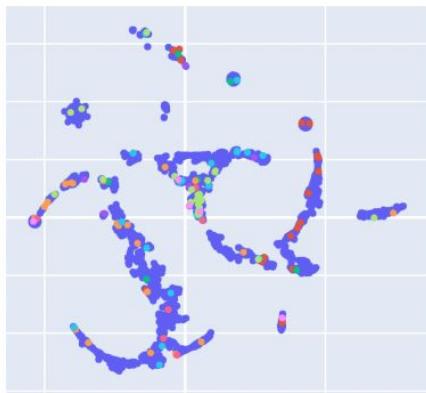
In the NG dataset we have **8 curated stories**. each story is linked to a unique set of paintings.

1. Women's Lives,
2. Contemporary Style and Fashion,
3. Water,
4. Women Artists and Famous Women,
5. Monsters and Demons,
6. Migration: Journeys and Exile,
7. Death, Battles and Commanders,
8. Warfare.

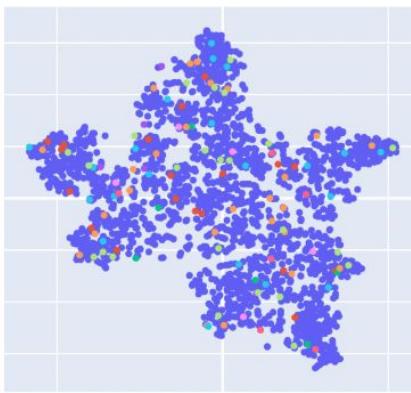
# Multi-Stakeholder aware RecSys



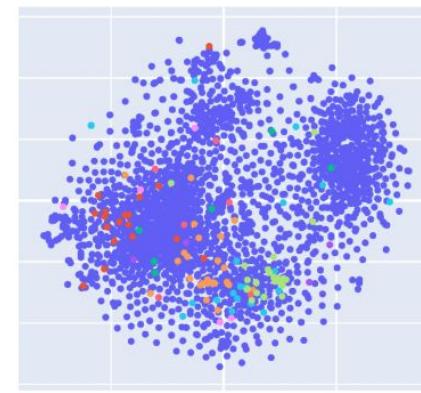
BERT



LDA



ResNet



## Story groups

- Uncategorised
- Water
- Migration\_Journeys\_and\_Exile
- Battles\_and\_Commanders
- Monsters\_and\_Demons
- Contemporary\_Style\_and\_Fashion
- Death
- Womens Lives
- Warfare

Latent space projection (t-SNE) of the curated story groups

## 1. POI recommendation →

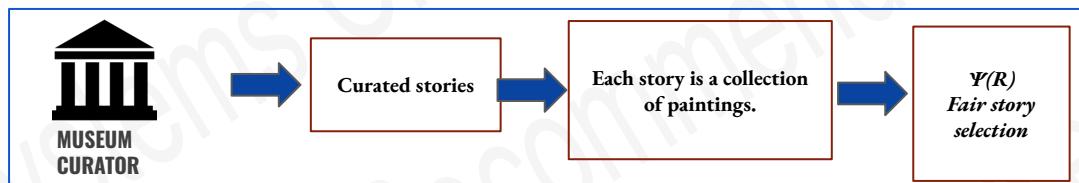
### 1.2 Matching Curator's Goal



- Increasing the number curated stories in the recommendation means fairly selecting the paintings from each story.
- We define a fair story selection function  $\Psi(R)$ .
- The function  $\Psi(R)$  rewards a typical diversity of stories in the recommendation set.

$$\Psi(R) = \sum_{i=1}^K \sqrt{\sum_{p \in S_i \cap R} \gamma_p}$$

- $S_i, i = 1, \dots, K$  is the story-partition of the dataset.
- $R$  is the recommendation set
- $\gamma_p$  is a representativeness score of story group carried by painting  $p$  in the recommendation set.

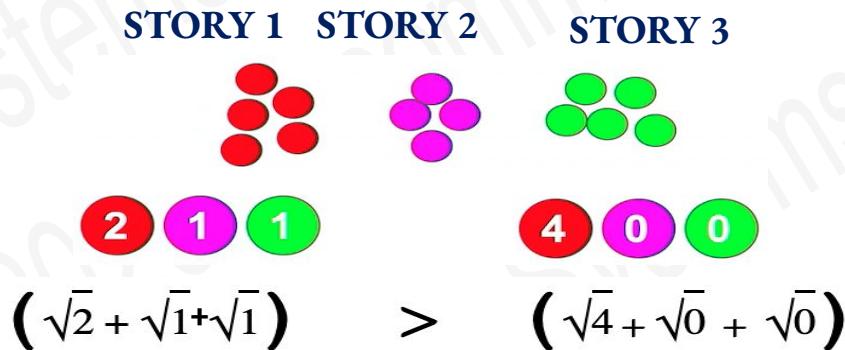


## 1. POI recommendation →

### 1.2 Matching Curator's Goal



$$\Psi(R) = \sum_{i=1}^K \sqrt{\sum_{p \in Si \cap R} \gamma_p}$$



*Representative & Informative Query Selection Yilmaz et al. SIGIR2015.*

## 1. POI recommendation →

*Recommend a set  $R$  of  $r$  paintings*

$$R(u) =$$

**Policy 1:** Maximize User Preference score.

- $\text{argmax } \sum_{a=1}^R S(P_a)$

**Policy 2:** Maximize the number of Curated stories.

- $\text{argmax } \Psi(R) = \text{argmax }$

$$\left( \sum_{i=1}^K \sqrt{\sum_{p \in Si \cap R} \gamma_p} \right)$$

$R(u)$

Respect time  
constraint

- $\text{argmax} \left( 1 - \varepsilon \left( \sum_{a=1}^R S(P_a) \right) + \varepsilon \left( \sum_{i=1}^K \sqrt{\sum_{p \in Si \cap R} \gamma_p} \right) \right)$

S.t

$$\sum_{a=1}^R T_v(P_a) \leq T_{ava}$$

User's tolerance  
to diversity

## 2. Path recommendation

We now have a recommendation Set of Paintings R.

- Project the painting from R on the Venues (Rooms).

The initial **Optimal route** should lead the visitor in the *least expensive path* and *highest relevance*.

Such that:

- *The total sum of estimated visiting times and travel time should not exceed the available time of the visitor.*
- *The crowd size in the selected rooms should not exceed the crowd tolerance threshold of the visitor.*



## 2. Path recommendation

Depending on the user **Relevance** could mean two things:

1. Quality: Visit the most interesting paintings.

- We define a quality score  $\Theta(v_i)$  which is the sum of the scores of all the recommended paintings in venue i.

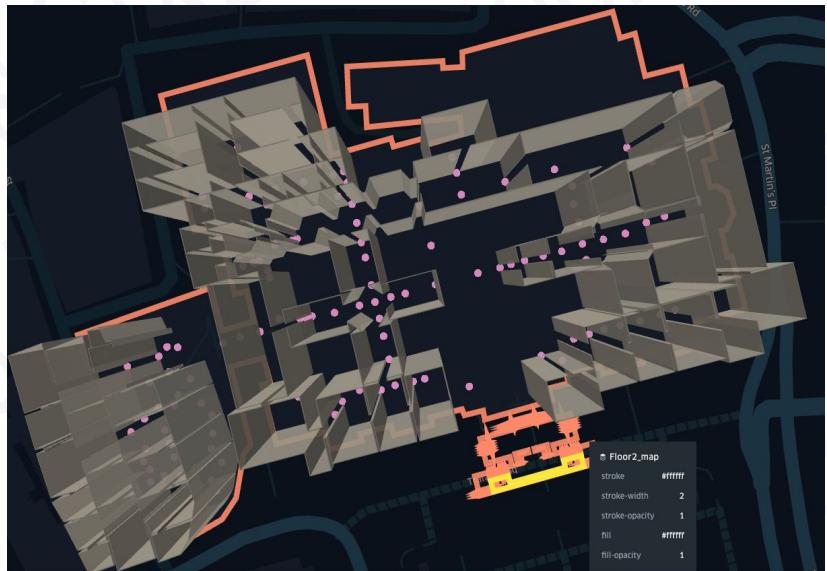
$$\bullet \quad \Theta(v_i) = \sum_{i=1}^h S(P_i)$$

- $h$  is the total number of recommended paintings in Venue V

2. Quantity: Visit as many paintings as possible.

- We define a quantity score of every venue

$$\delta(v_i) = b_i$$





## 2. Path recommendation

*Recommend a path  $PT(u)$   
(sequence of  $M$  rooms)*

**Policy 1:** Maximize relevance score  $S(R)$

- $\text{argmax } \sum_{a=1}^M \Theta(v_a) \quad \text{if Quality} > \text{Quantity}$
- $\text{argmax } \sum_{a=1}^M \delta(v_a) \quad \text{otherwise}$

**Policy 2:** Minimize travel distance

$$\bullet \text{ argmin } \sum_{a=1}^M \text{dist}(v_a, v_{a+1})$$

**Subject to:**

1. Time constraint:

$$\bullet \sum_{a=1}^M T(v_a) + Tt \leq T_{ava}$$

$$Tt = \sum_{a=1}^M Tt(v_a, v_{a+1})$$

2. Crowd constraint:

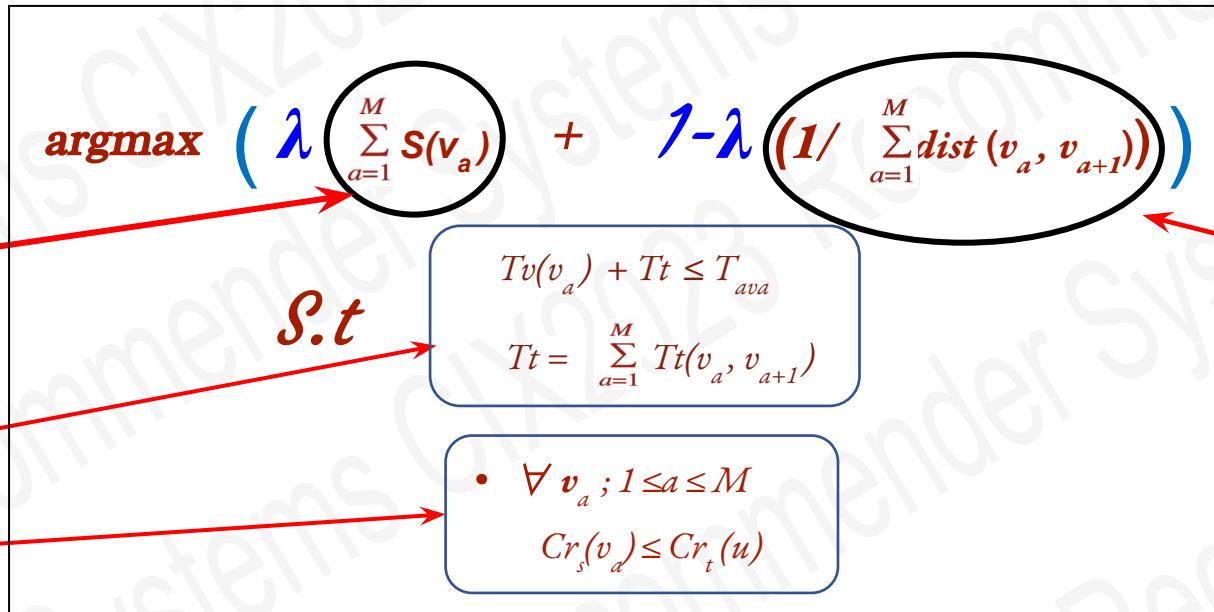
$$\bullet \forall v_a ; 1 \leq a \leq M$$

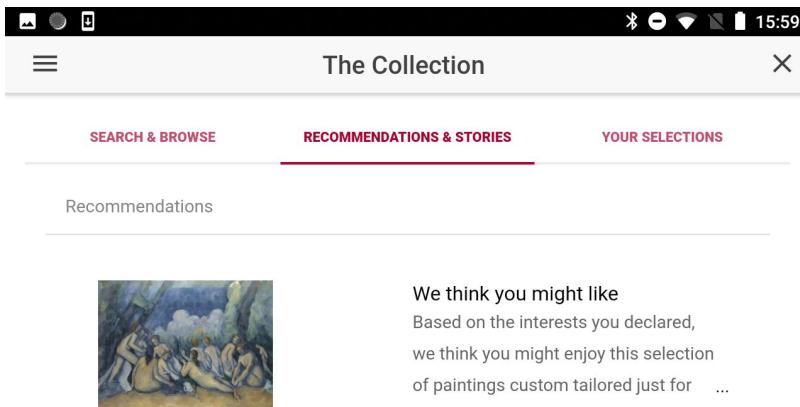
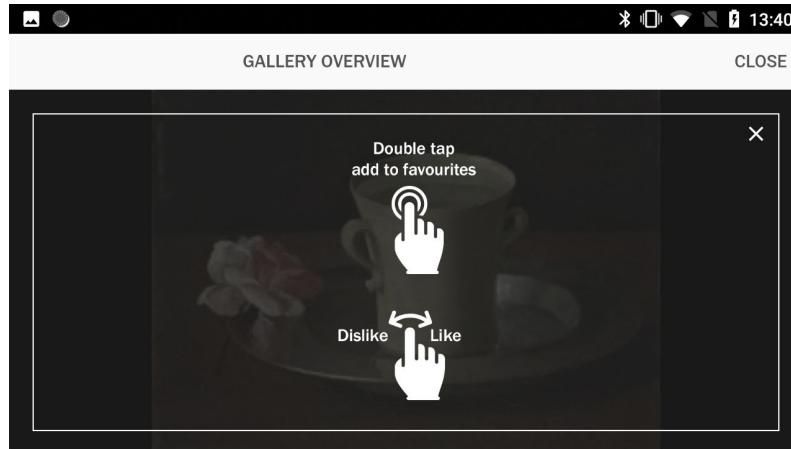
$$Cr_s(v_a) \leq Cr_t(u)$$

## 2. Path recommendation



*Recommend a path  $PT(u)$   
(sequence of  $M$  rooms)*





Personalized ? popularity, curator's theme,

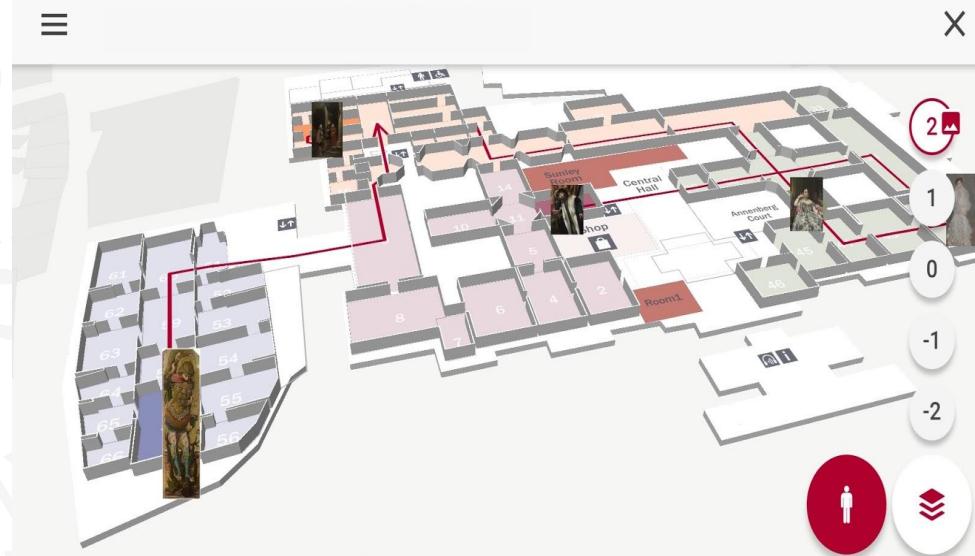
# Multi-Stakeholder aware RecSys



Mobile app

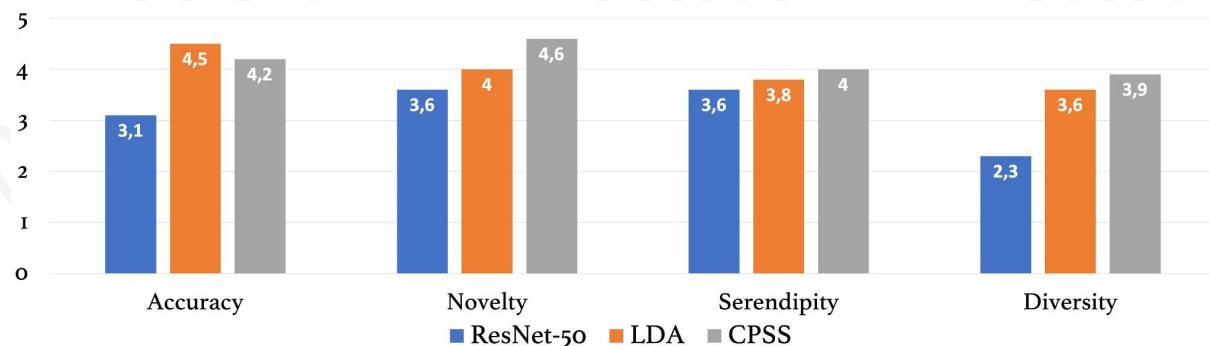


Mixed reality glass



# Multi-Stakeholder aware RecSys

**Baseline** Single objective **Vs** Multi-objective



Yilma et al. (UMAP '21)

# Multi-Stakeholder aware RecSys



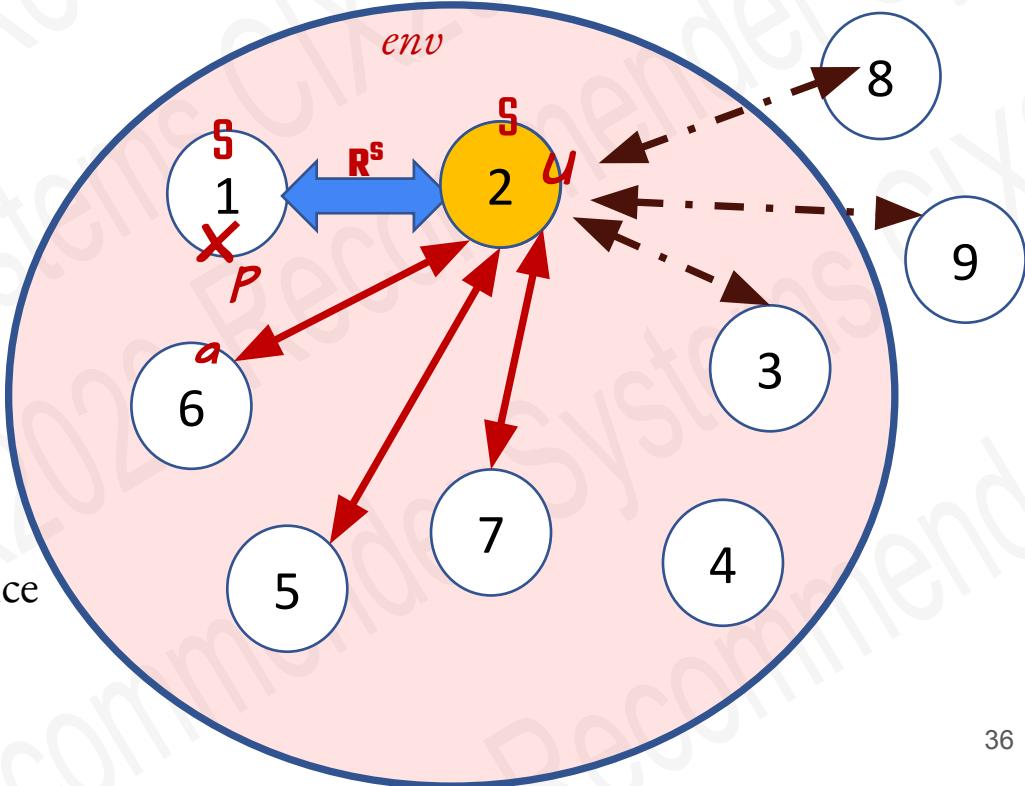
- Understanding the Problem setting.
- Identifying stakeholders (**Competing objective or Constraints**).
- Prioritizing objectives.

# Formulating a RecSys Problem



- Smart system environment  $env$ ,
- **Personalisation** is a function of a social component **S** of a system.
- **Personaliser**( $x_{pa}$ );
- **User**( $U$ )
- **Crowd**( $Cr$ ): direct influence
- **Context elements**( $Cx$ ): indirect influence

$$Pa^{coss} = f(u, x_{pa}, cr, cx, env)$$





# Modern RecSys paradigms

## What are LLMs?

LLMs are machine learning models capable of Natural Language Processing (NLP), as they are trained on huge amounts of text data (usually from the internet/books) via deep-learning algorithms.

- **Pre-trained language models:** encode an extensive amount of world knowledge, and they can be applied to a multitude of downstream NLP applications with **zero** or **just a handful of demonstrations**.

# Large Language Models (LLMs) as RecSys



User Preference

Pre-trained  
LLM

Recommendations

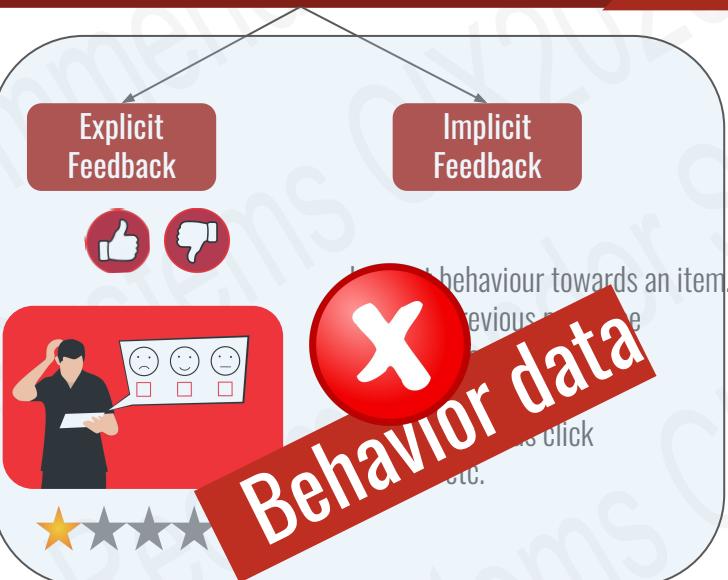
Explicit  
Feedback



Implicit  
Feedback



X  
**Behavior data**



Store knowledge that can complement the behavior data



Predict future behaviours

buying turkey on Thanksgiving day

# Large Language Models (LLMs) as RecSys



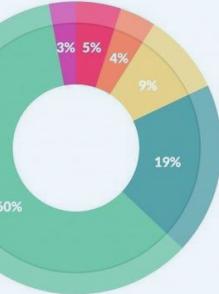
## Data Pre-processing



## Pre-trained LLM

## Post Processing

## Evaluation



What data scientists spend the most time doing

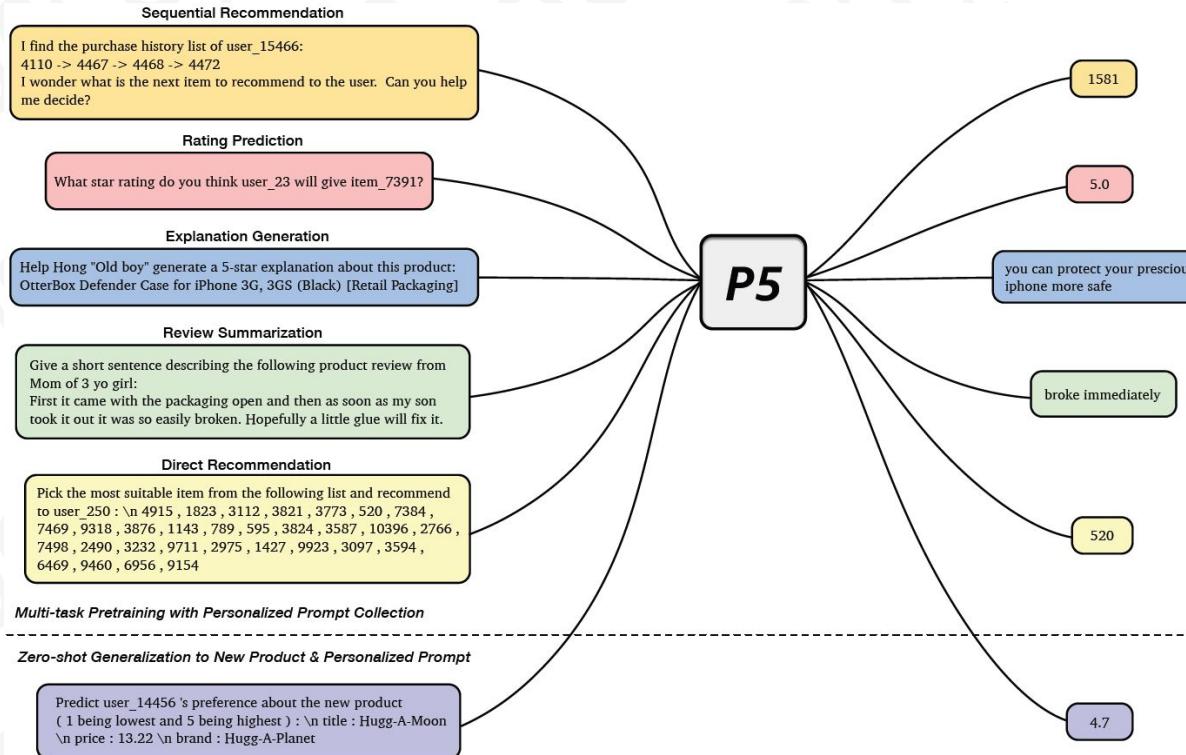
- Building training sets: 3%
- Cleaning and organizing data: 60%
- Collecting data sets: 19%
- Mining data for patterns: 9%
- Refining algorithms: 4%
- Other: 5%



- Sort
- Filter
- Recommend



## 1. Pretrain, Personalized Prompt & Predict Paradigm (P5)



## 2. M6-Rec: Generative Pretrained Language Models are Open-Ended Recommender Systems.

*Cui et al. 2022*

- Model pre-trained using Alibaba's LLM (M6)
- Convert recommendation tasks
  - Retrieval,
  - Scoring,
  - Explanation generation
  - Conversational recommendation

### prompt

For zero-shot scoring tasks



or



- 
- 

**Language understanding or Generation**

[BOS'] A user clicks hiking shoes [EOS'] [BOS] also clicks trekking poles [EOS]  
 [BOS'] A user clicks hiking shoes [EOS'] [BOS] also clicks yoga knee pads [EOS]

user-side features

user + item related features

The plausibility of each is calculated based on token probability from M6.

Avoids using user IDs or item IDs as input

## Prompt-Tuning Pretrained Language Models

### 3. Language Model Recommender Systems (LMRecSys)

Zhang et al. 2022

- Uses the user's past interaction history (watched movies), taken from the MovieLens-1M dataset, to predict the next movie that the user would watch.

 prompt

A user watched One Flew Over the Cuckoo's Nest, James and the Giant Peach, and My Fair Lady. Now the user wants to watch [MASK].



Estimate the probability distribution of the next item

## Prompt-Tuning Pretrained Language Models

4. Zero-Shot Next-Item Recommendation (NIR) *Wang et al. 2023*
5. Augmented Recommender System (Chat-REC) *Gao et al. 2023*



# Pros & Cons of using LLMs as RecSys



## Pros

- Zero-shot and few-shot capabilities
  - Sparsity scenarios (**cold start, long tail**)
- Adapting to new information
  - Without changing architecture/ retraining
- Conversational- user engagement
- No need for large amount of data
- Generalization capability: not task specific
- **Explanation Generation**

## Cons

- LLMs are large, complex, and costly to train
- Highly sensitive to input prompts and may not fully follow the instruction prompt.
- Recommendations can be highly **dataset** dependent.
- LLMs may recommend items that are not present in the candidate set.

*Zhang et al. 2023, suggested feeding the candidate item set into prompt.*



# Pros & Cons of using LLMs as RecSys

## Case 1

You are a movie recommender system now.

*\{\{Examples\}\}}*

Input: Here is the watching history of a user: Aliens, E.T. the Extra-Terrestrial, Contact, The Matrix, X-Men. Based on this history, would this user prefer The Fox and the Hound or Steamboat Willie? Answer Choices: (A) The Fox and the Hound (B) Steamboat Willie

Output: The answer index is N/A as neither option is relevant to the user's watching history.

## Case 2

You are a book recommender system now.

*\{\{Examples\}\}}*

Input: Here is the reading history of a user: The Cellist of Sarajevo, After I'm Gone: A Novel, The Reason I Jump: The Inner Voice of a Thirteen-Year-Old Boy with Autism, The Serpent of Venice: A Novel, We Are All Completely Beside Ourselves: A Novel. Based on this history, would this user prefer The Secret Life of Bees or The Help? Answer Choices: (A) The Secret Life of Bees (B) The Help

Output: The answer index is N/A. It is difficult to determine the user's preference based on this reading history as neither book is similar in genre or theme to the books they have read.





# Common Issues and Challenges in RecSys

# Multi-Stakeholder aware RecSys



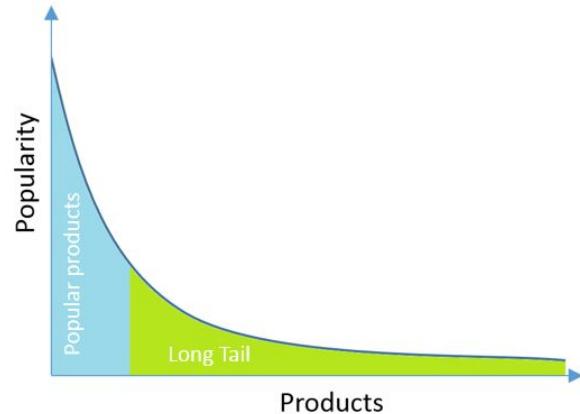
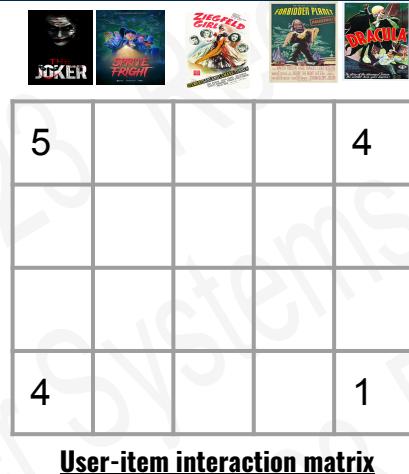
- Understanding the Problem setting.
- Identifying stakeholders (**Competing objective or Constraints**).
- Prioritizing objectives.

# RecSys Issues & Challenges

- **Cold start**

→ Sparsity

$$\text{Sparsity} = \frac{\# \text{ ratings}}{\text{total } \# \text{ cells}}$$



- **Diversity**

# RecSys Issues & Challenges

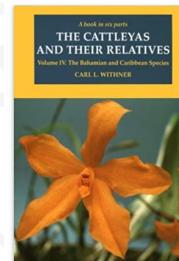
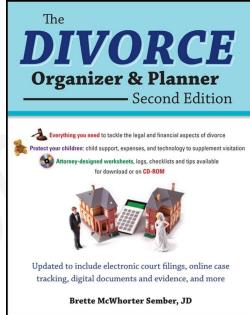
- **Privacy** : Third party, sensitive private information
  - **Collaborative Filtering**: users willingly disclose their preferences to the system in the hope of getting useful recommendations.



Likes growing Orchid



Bought



Customers who viewed this item also bought

The Divorce Organizer and Planner with CD-ROM, 2nd Edition  
by Brette Sember

10 offers from \$5.13

4.5 stars, 10 reviews

Paperback  
\$7.99  
\$16.63 shipping

CHILD CUSTODY JOURNAL

Child Custody Battle & Co-parenting Log Book, Detailed Record Journal...

Annie Hyatt

4.5 stars, 63 reviews

Paperback  
\$6.97  
\$16.68 shipping

CHILD CUSTODY LOG BOOK

Child Custody Battle With... Custody Battle With... Visitation Planner...

Eric Reeves Publication

4.5 stars, 22 reviews

Paperback  
\$6.99  
\$17.27 shipping

child custody log book

Child Custody Log Book: Child Support and Custody Journal Visitation Planner...

Eddie M Sanchez Publishing

4.5 stars, 23 reviews

Paperback  
\$6.99  
\$16.81 shipping

Child Custody Journal: Track Expenses, Communication, Visitation, Witness, an...

Samir Schoen

4.5 stars, 29 reviews

Paperback  
\$6.98  
\$18.87  
\$16.73 shipping

CO-PARENTING WITH A NARCISSISTIC EX: How to Protect Your Child From a Toxic Parent &...

Lara Carter

4.5 stars, 23 reviews

Paperback  
\$6.99  
\$16.81 shipping

Child Custody Journal: Child Custody Visitation Planner Calendar and Co-Parenting Log Book to Plan and Record...

Charlie L. Press

4.5 stars, 29 reviews

Paperback  
\$6.99  
\$16.81 shipping

See all formats and editions

The Cattleyas and Their Relatives: Volume IV: The Bahamian and Caribbean Species Hardcover – April 15, 1996

by Carl L. Witthner (Author)

4.6 stars, 3 ratings

Book 4 of 5: Cattleyas and Their Relatives





# RecSys Issues & Challenges



## Privacy

- Complete privacy may not be realistic → compromise on minimizing the privacy breaches
- Comes at the expense of the accuracy of the recommendations.
- Important to analyze this trade-off carefully.



# RecSys Issues & Challenges



- **Adaptivity:** changing business needs
- **Robustness:** Attack/ stress
- **Scalability**

# RecSys Issues & Challenges

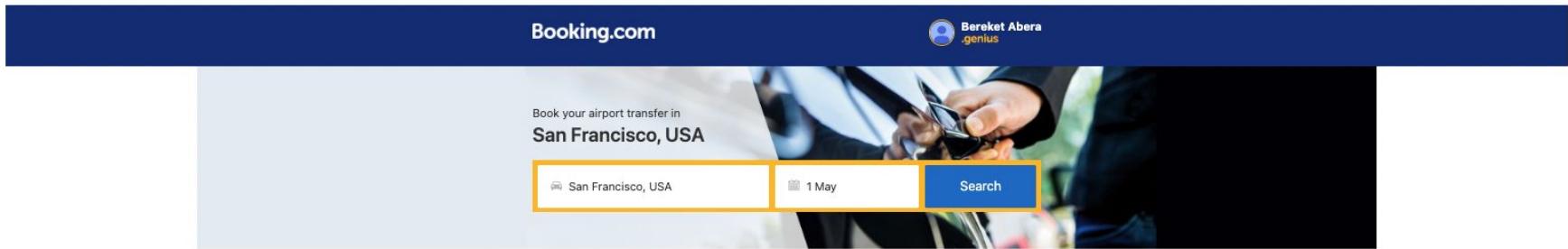
- **Proactiveness:** Predict when and how to push Recommendations ← implicit request

Bereket, book a taxi direct to Holiday Inn Express Hotel & Suites Fisherman's Wharf ➔ [Inbox](#)



Booking.com <[email.campaign@sg.booking.com](mailto:email.campaign@sg.booking.com)> [Unsubscribe](#)  
to me ➔

Apr 27, 2019, 11:55 PM [Star](#) [Reply](#) [Forward](#)



The screenshot shows an email from Booking.com. At the top, there's a dark blue header bar with the Booking.com logo on the left and a profile picture of 'Bereket Abera .genius' on the right. Below the header, the main content area has a white background. On the left, there's a large image of two people in a car. Overlaid on the image is a white box containing text: 'Book your airport transfer in San Francisco, USA'. Below this text is a search bar with three input fields: 'San Francisco, USA' (with a car icon), '1 May' (with a calendar icon), and a blue 'Search' button. The rest of the email body is mostly cut off by the bottom of the screenshot.

Hi Bereket,

Looking for a convenient, reliable way to get to your accommodation? Your driver and car will be waiting at the airport – no unfamiliar public transport to navigate, no worries on finding the way to your stay.

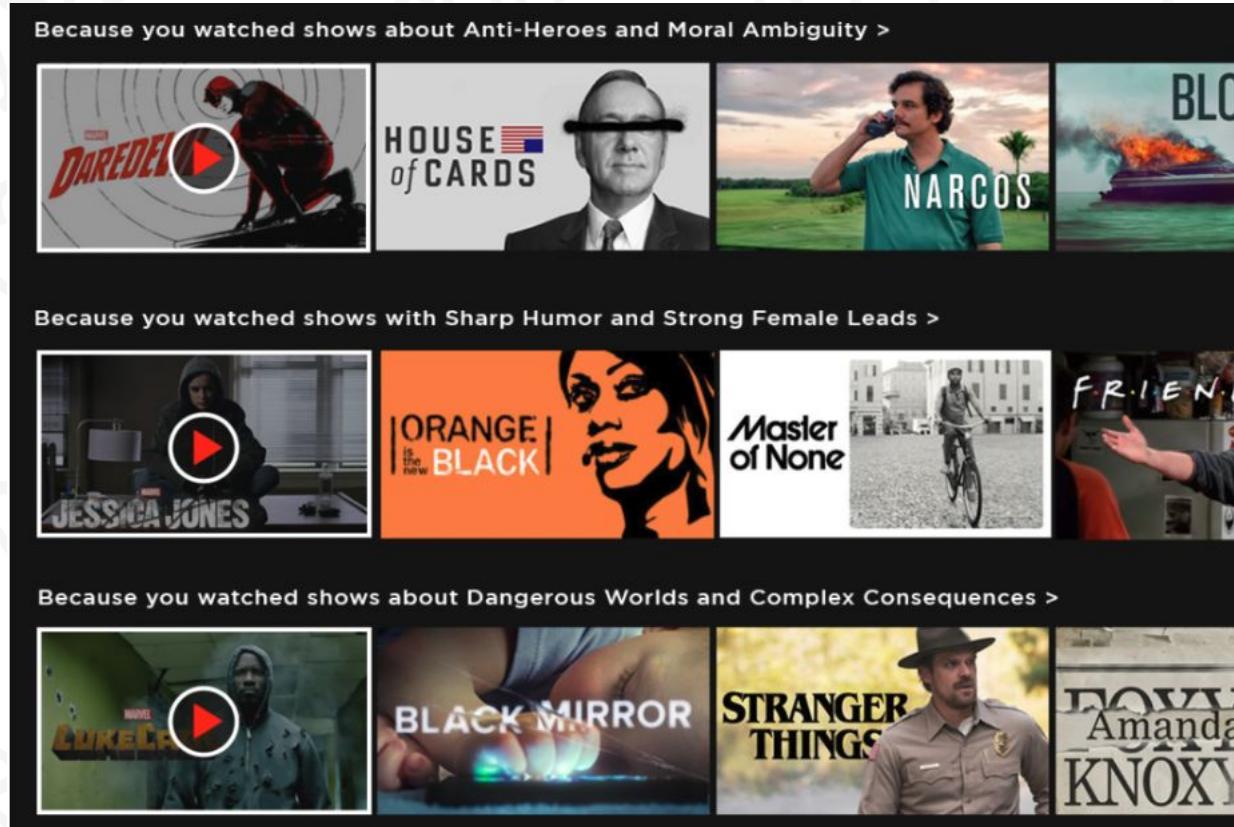
✓ **Guaranteed price**   ✓ **Guaranteed pick-up**   ✓ **Tried and trusted drivers**

Want to get back to the airport just as easily? Book your return journey at the same time and enjoy your holiday with one less thing to worry about.

[Find your ride](#)

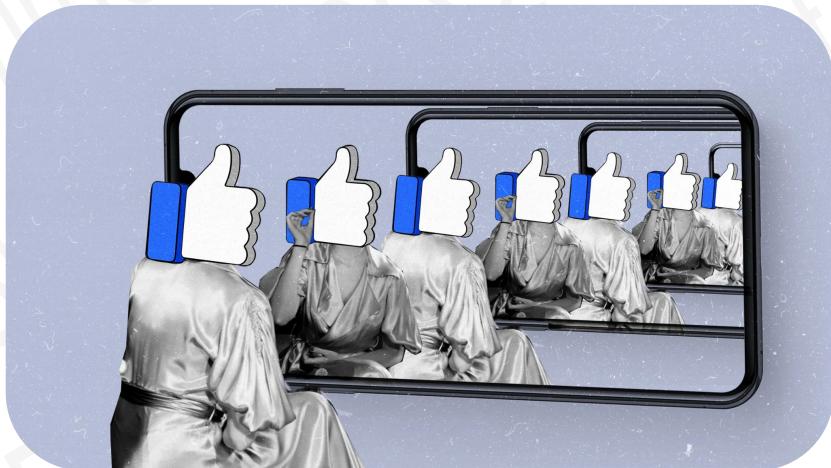
# RecSys Issues & Challenges

- Explainability



# RecSys Issues & Challenges

- Echo Chamber



# RecSys Issues & Challenges

- **Filter bubbles**



- Limiting exposure to diverse perspectives.
- Reinforcing existing biases and stereotypes.

# Course Reflection



Contact: [bereket.yilma@uni.lu](mailto:bereket.yilma@uni.lu)



[Feedback form](#)



# Come to Luxembourg for research

- Research visit
- Internship
- PhD
- Postdoctoral fellowship



## Computational Interaction (COIN) Lab

- **Human-Computer Interaction**
  - Emotion recognition using EEG, FNIRS, Eye tracking
  - Brain computer interfaces (BCI)
  - Bio-signal processing
  - Mouse tracking
  - Adaptive User Interfaces (AUI)
- **Machine Learning (Applied)**
- **Personalisation, RecSys**
- **Conversational Agents, Speech synthesis**

[Why should you come to Luxembourg for research?](#)



The 1<sup>st</sup> International Workshop on  
Human-in-the-Loop Applied Machine Learning  
(HITLAML)



# CALL FOR PAPERS

## MAIN DEADLINES (AOE)

### PAPER & DEMO SUBMISSION DEADLINE

~~June 15, 2023 (AoE)~~

**July 14, 2023 (AoE)**

### TRAVEL GRANT APPLICATION DEADLINE:

**July 20, 2023 (AoE)**

### NOTIFICATION OF ACCEPTANCE

**July 31, 2023**

### REGISTRATION OPENS

**August 1, 2023 (#Free of Charge)**

### WORKSHOP DATE

**September 04 - 06, 2023**



**BELVAL, LUXEMBOURG**

<https://hitlaml.uni.lu>

