Recommender Systems

Invited talk at

Elements of AI

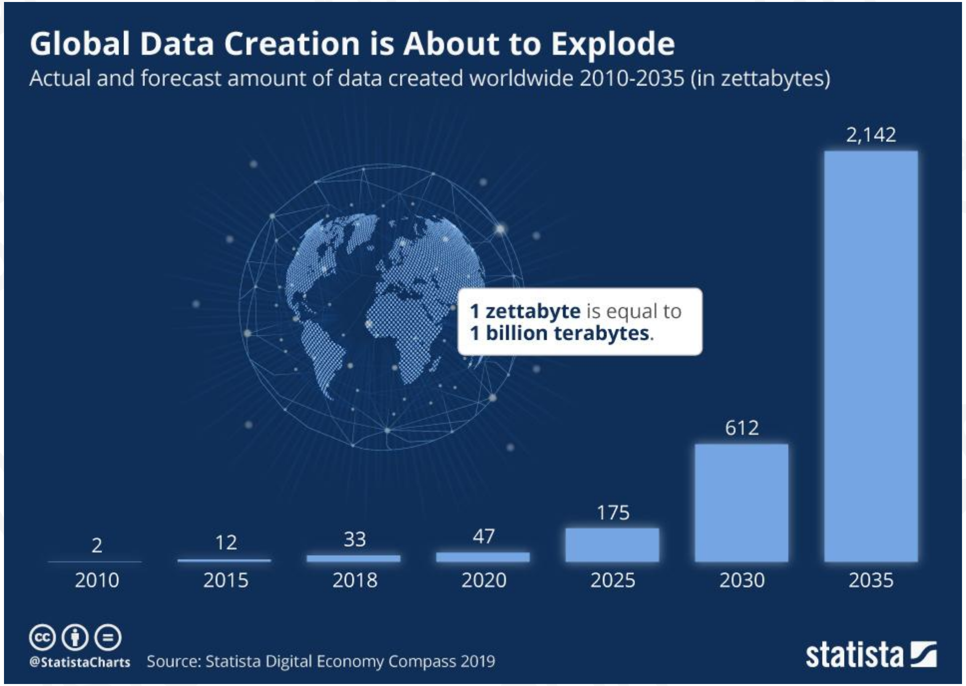
Dr. Bereket A. Yilma

Outline

- ➔ Background
    - ◆ Why do we need RecSys?
    - ◆ Real World examples
- ➔ RecSys Paradigms
- ➔ How to apply RecSys
    - ◆ The RecSys Pipeline a case-study approach
- ➔ Open challenges.

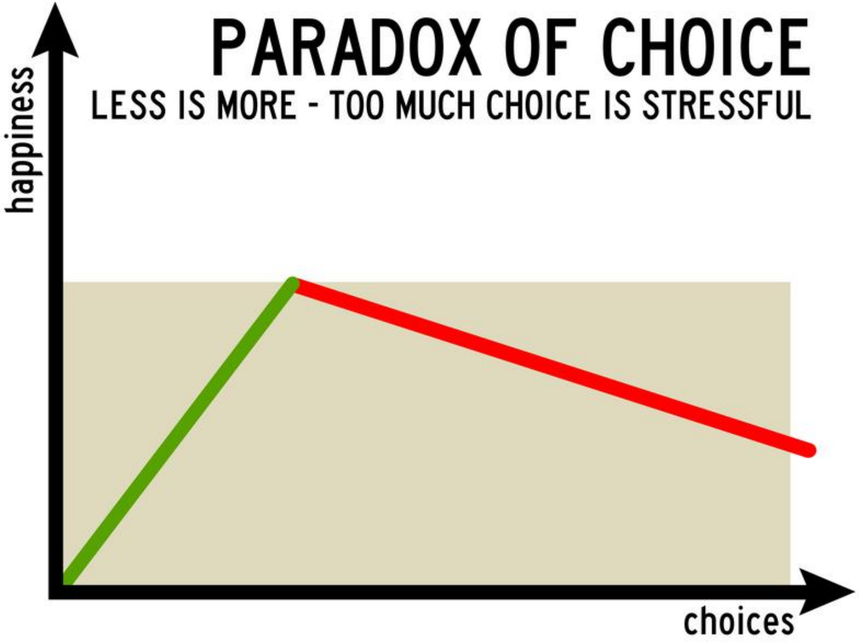
Background

Our options on the web are limitless



Background

Abundance creates a problem



Background

UNIVERSITÉ DU LUXEMBOURG

Source: Politiken (Based on Our Yale List)

Decision making has become extremely challenging with the overwhelming number of products and services.

Background

Abundance of Options
Certainty of Decisions
RECOMMENDER SYSTEMS

Recommender Systems (RecSys)

- “Algorithms that provide suggestions for items that are **presumed** most pertinent to a particular user.” [1]

- Typically, the suggestions refer to various decision-making processes such as:

80%
of content consumed on Netflix is due to recommendations.

60%
of video clicks on Youtube’s homepage are attributed to recommendations

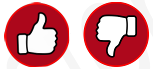
35%
of its revenue is generated by its recommendation engine

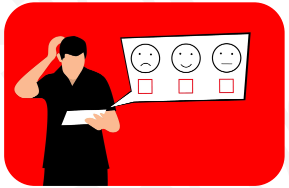
  - ➢ What product to purchase,
  - ➢ What music to listen to,
  - ➢ Which movies to watch,
  - ➢ What news to read,
  - ➢ Which route to take,
  - ➢ which place to visit, etc.

Recommender Systems (RecSys)

User Preference → Recommender System → Recommendations

- Explicit feedback
- Implicit feedback





Indirect behaviour towards an item.

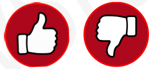
- previous purchase
- previous watch
- previous read
- previous click
- etc.

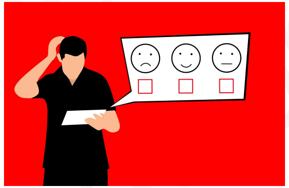
Predict future behaviours

2. Recommender Systems (RecSys)

Data	Machine Learning Model	Predictions
User Preference	Recommender System	Recommendations

Explicit feedback





Implicit feedback

Indirect behaviour towards an item.

- previous purchase
- previous watch
- previous read
- previous click
- etc.

Predict future behaviours

RecSys Paradigms

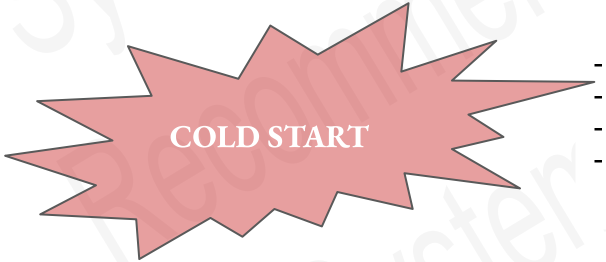
- There are three common approaches namely Collaborative Filtering, Content-Based filtering and Hybrid Recommender Systems.

1. Collaborative Filtering

"Similar users like similar things"

Users

- Subscribers
- Readers
- Buyers
- etc.



New item

5	?	1	?	4
?	?	?	?	?
2	?	?	3	?
4	?	2	?	1

User-item interaction matrix

- Ratings
- Time spent
- Product clicked

"Customers who bought this item also bought"

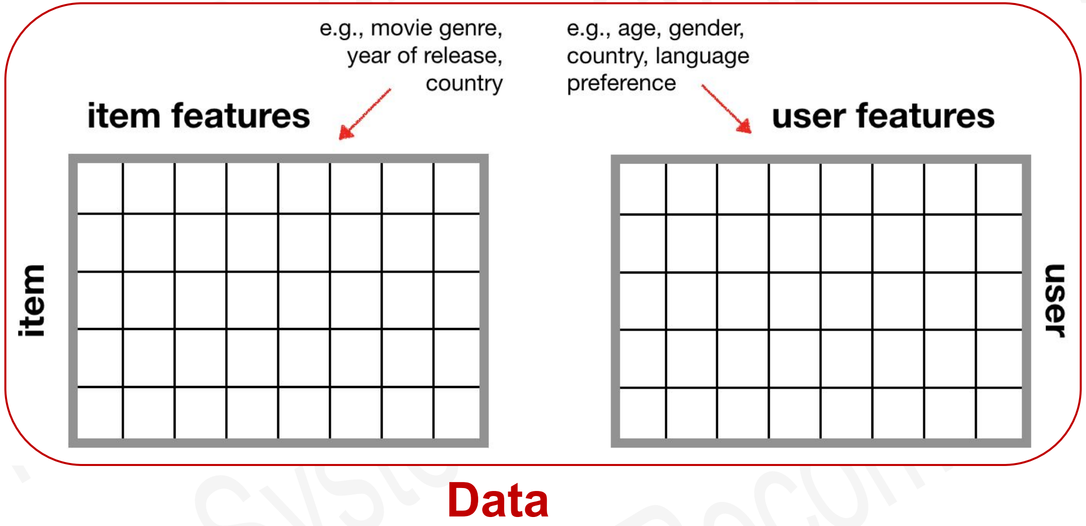
New user

Data

RecSys Paradigms

2. Content-Based Filtering:

- Works based on the comparison of the analogy between the **user's profile** and **content of the items**.
- Use additional information about **users** and/or **items (“Features”)** that explain the observed **user-item** interactions.



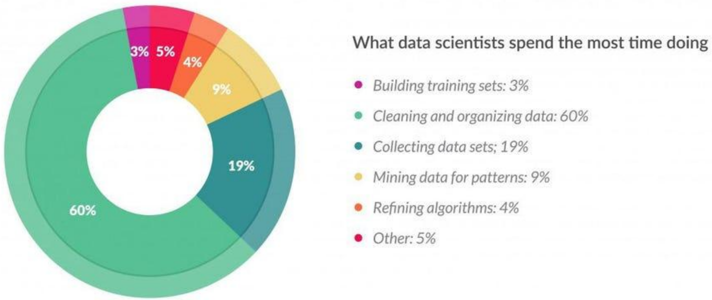
- Suffer less from cold start problem.
- only new users/items with unseen features suffer from cold start problem.

3. Hybrid Recommender Systems:

- Combine CF and CBF approaches
- Usually take two forms
  - Train two models independently (one CF model and one CBF model) and combine their suggestions.
  - Directly build a single model that unifies both approaches (often a Neural Network)
    - Input (Prior information user/item) + user-item interaction

The typical RecSys Pipeline

Data Pre-processing



Model Training

Post Processing

Evaluation

- Sort
- Filter
- Recommend

The RecSys pipeline: A case-study approach

The RecSys pipeline: A case-study approach

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

The RecSys pipeline: A case-study approach

DISCLAIMER

CROSS CULT

LIST

European Commission

HORIZON 2020

The RecSys pipeline: A case-study approach

Personalised Visual Art Recommendation

Context: National Gallery, London

- >= 2,300 paintings dating from the mid-13th century to 1900.
- Total floor area of 46,369 square meters, 3 floors.
- 6.2 million visitors/year (2019)

The RecSys pipeline: A case-study approach

Task — **Personalised Visual Art Recommendation**

Data — **Visual Art (Paintings)**

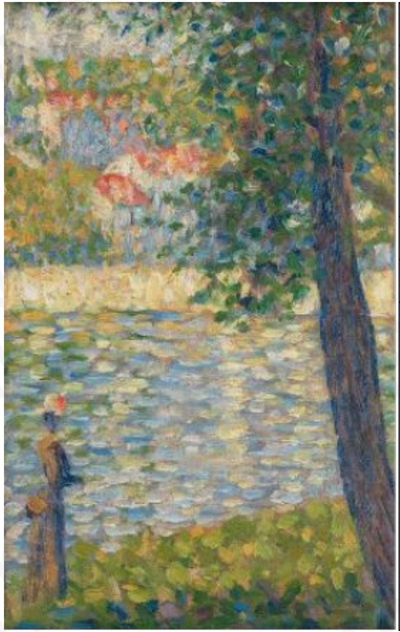
Target User — **Visitors (New)**

***assumptions Visitor profile**

The RecSys pipeline: A case-study approach

1. Data: Visual Art (Paintings)

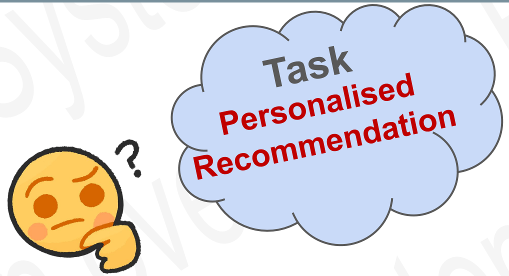
- 2,368 painting

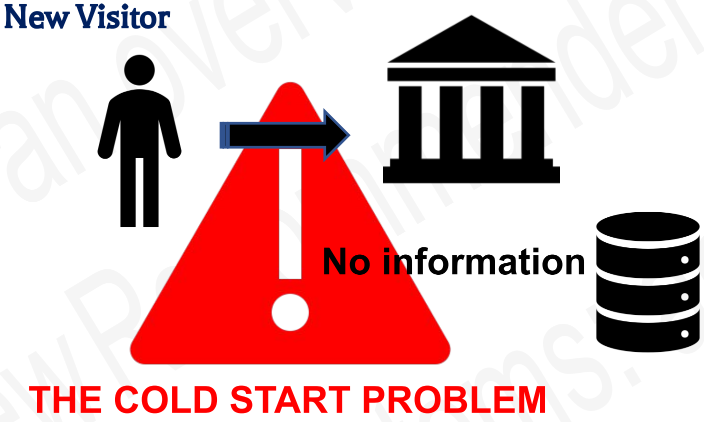


painting_id	000-018P-0000
title	The Morning Walk
artist	Georges Seurat
publication_date	19th_century
size_format	Portrait
size	Very Small
technique	oil painting
description	A woman, silhouetted against the shimmering water, strolls along a riverbank. The red roofs of houses can be made out along the opposite bank. Between 1882 and 1886 Seurat painted numerous such landscape studies on small wooden panels, some as independent works and others in preparation for his large-scale compositions. This sketch provided the starting point for a painting of 1885, 'The Seine at Courbevoie' (private collection).

The RecSys pipeline: A case-study approach

Data Pre-processing



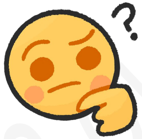


Query User (Profiling)

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

The RecSys pipeline: A case-study approach

Data Pre-processing



Task
Personalised Recommendation

Model Training

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

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	0.59
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
0.59	0.67	0.67	0.66	0.39	0.39	0.54	0.47	0.4	0.44	0.58	0.66	0.53	0.53	1	0.9
0.52	0.6	0.59	0.59	0.39	0.34	0.45	0.39	0.37	0.42	0.51	0.61	0.45	0.46	0.9	1

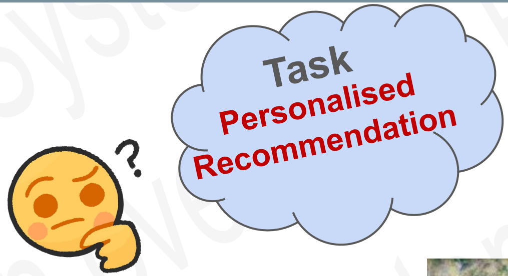
- If a user likes painting **A** find paintings **B, C, D** that are similar to A.

The RecSys pipeline: A case-study approach

RECOMMENDER SYSTEMS

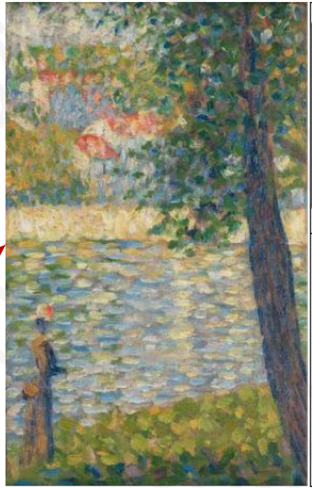
UNIVERSITÉ DU LUXEMBOURG

Data Pre-processing



Metadata

Textual description

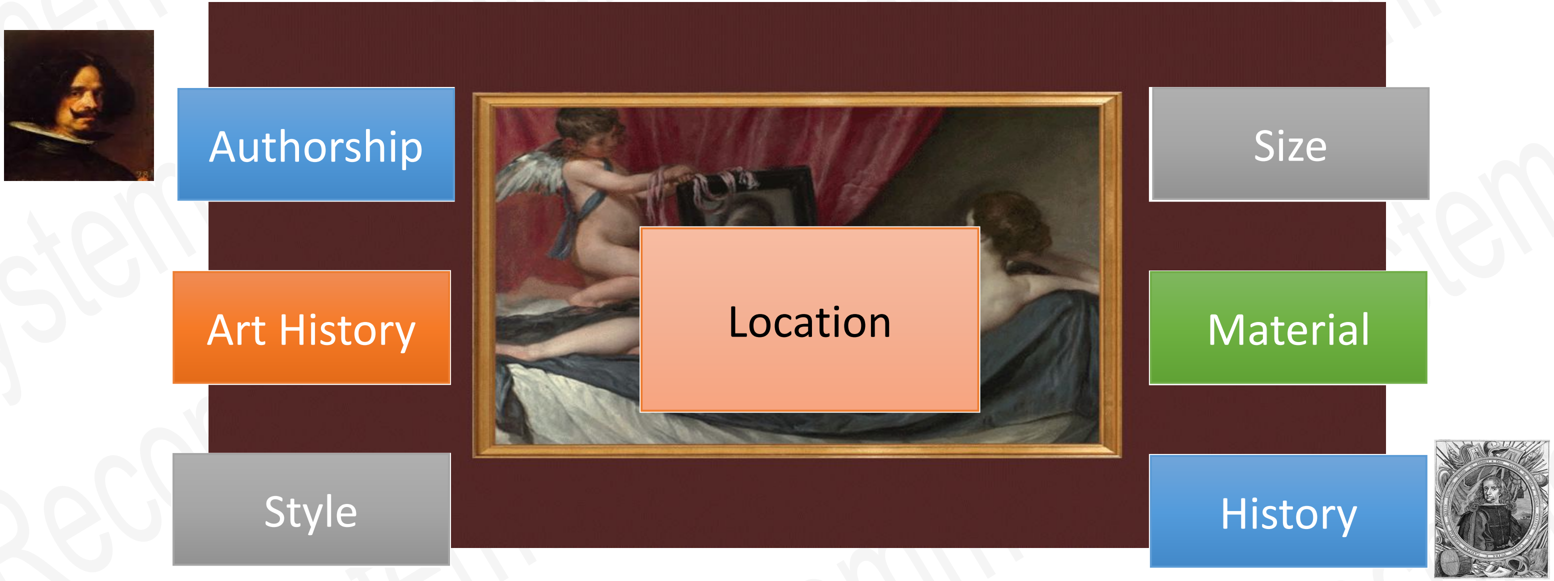


painting_id	000-018P-0000
title	The Morning Walk
artist	Georges Seurat
publication_date	19th_century
size_format	Portrait
size	Very Small
technique	oil painting
description	A woman, silhouetted against the shimmering water, strolls along a riverbank. The red roofs of houses can be made out along the opposite bank. Between 1882 and 1886 Seurat painted numerous such landscape studies on small wooden panels, some as independent works and others in preparation for his large-scale compositions. This sketch provided the starting point for a painting of 1885, 'The Seine at Courbevoie' (private collection).

Image

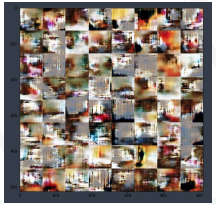
The RecSys pipeline: A case-study approach

Manually curated metadata to drive recommendations.



The RecSys pipeline: A case-study approach

Train a model that can learn Visual features from images of paintings



Some issues:

- Recommendations don’t have direct interpretation.
- Often fail to capture complex semantics such as triggered reflections

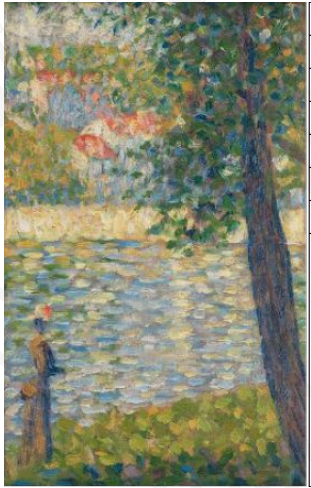
The RecSys pipeline: A case-study approach

RECOMMENDER SYSTEMS

UNIVERSITÉ DU LUXEMBOURG

Data Pre-processing → Task: Personalised Recommendation → Model Training

Textual description

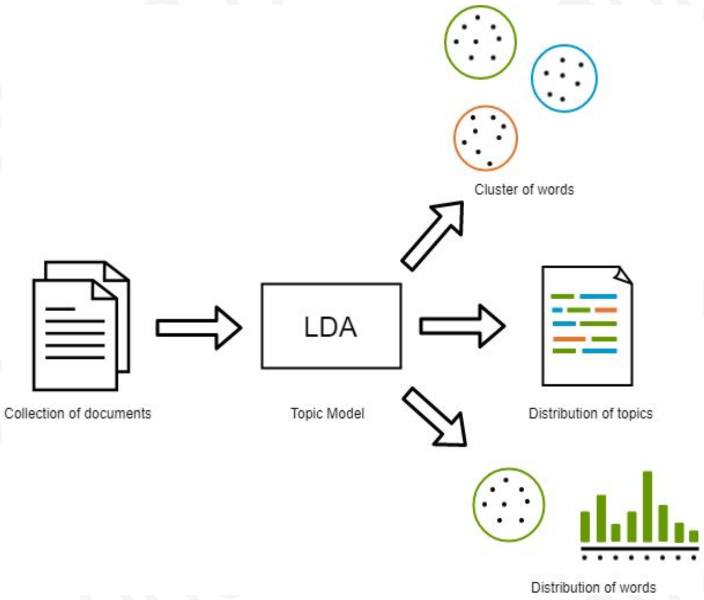


painting_id	000-018P-0000
title	The Morning Walk
artist	Georges Seurat
publication_date	19th_century
size_format	Portrait
size	Very Small
technique	oil painting
description	A woman, silhouetted against the shimmering water, strolls along a riverbank. The red roofs of houses can be made out along the opposite bank. Between 1882 and 1886 Seurat painted numerous such landscape studies on small wooden panels, some as independent works and others in preparation for his large-scale compositions. This sketch provided the starting point for a painting of 1885, 'The Seine at Courbevoie' (private collection).

Topic Modelling

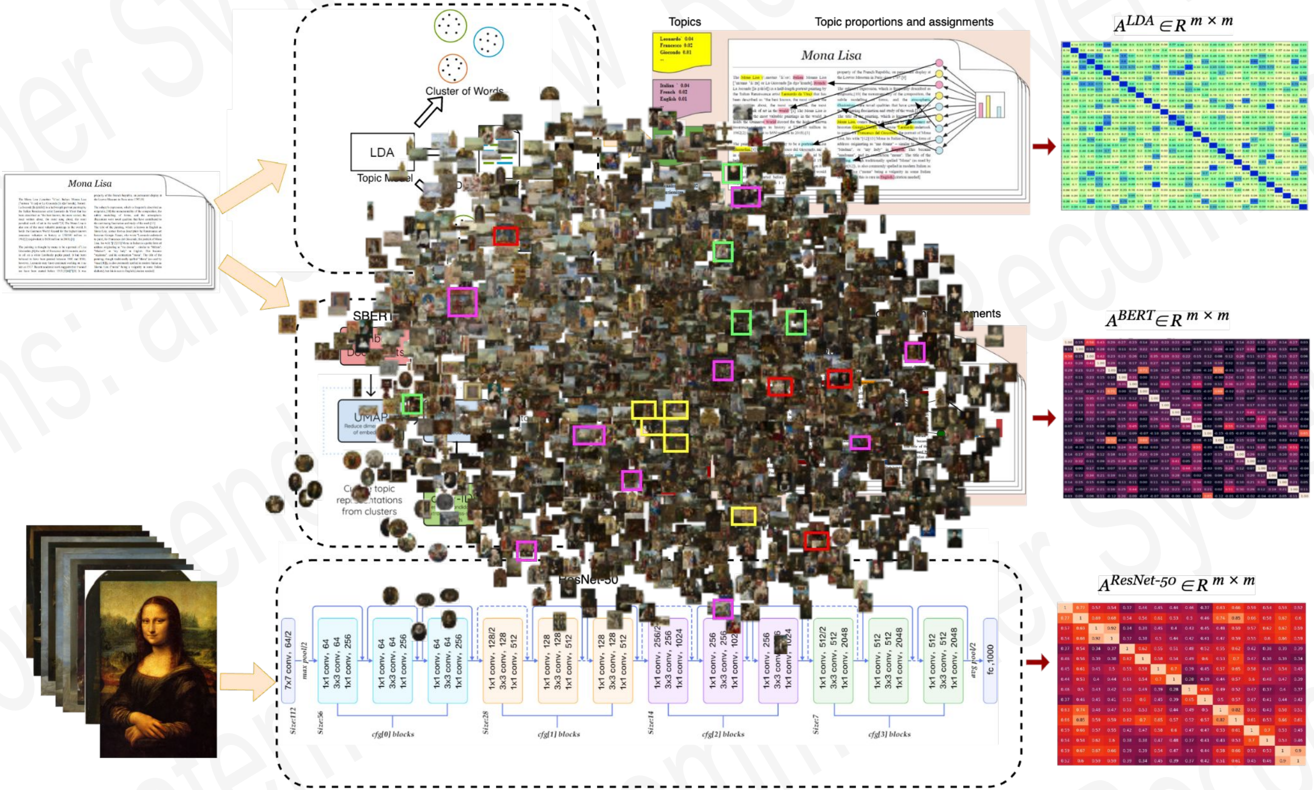
- LDA is an unsupervised generative probabilistic model.

- Particularly it can identify similar documents by uncovering abstract topics that occur in a collection of documents.



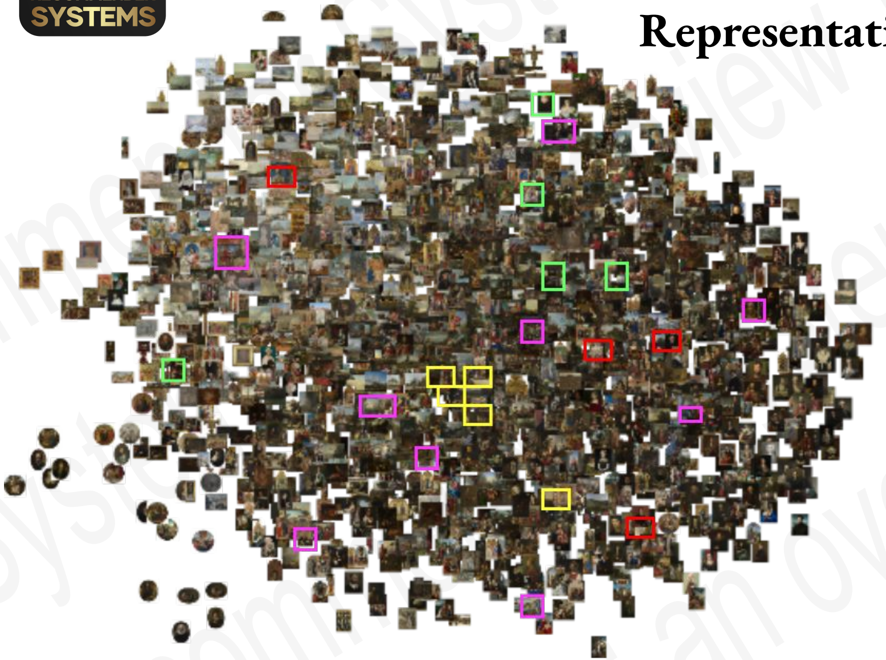
Yilma et al. 2020

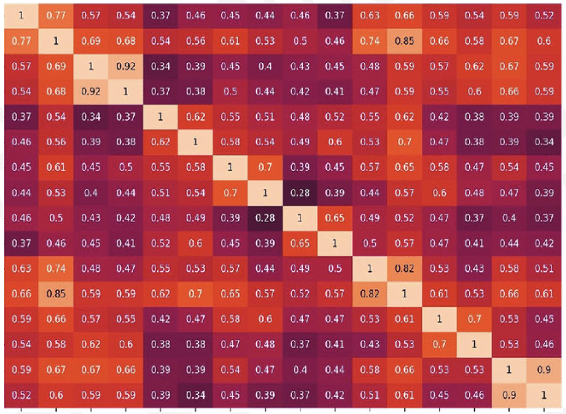
The RecSys pipeline: A case-study approach



The RecSys pipeline: A case-study approach

Representation learning:

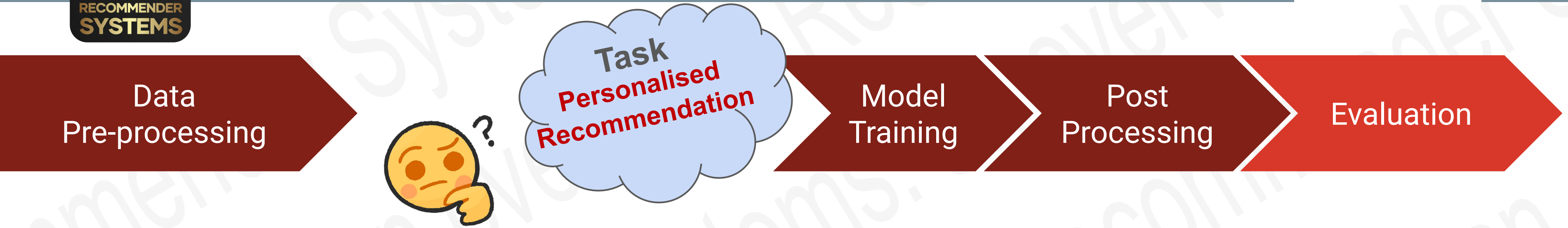




Similar paintings will be represented close to each other in the representation space

RecSys is all about efficient Search and Retrieval

The RecSys pipeline: A case-study approach



- Proxy for how relevant/ good are recommender systems.

1. Offline Experiment.

- Easiest to conduct
- Requires no interaction with users.

2. User Studies

- Small group of Subjects
- Controlled setting
- Qualitative/quantitative

3. Online Experiments

- The most trustworthy
- A pool of real users (unaware of the experiment)

- Common evaluation protocols in ML, IR

RECOMMENDER SYSTEMS

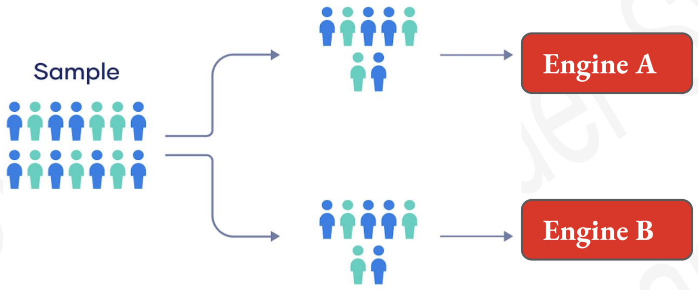
UNIVERSITÉ DU LUXEMBOURG

Evaluation: User Studies

Between vs. Within Subjects

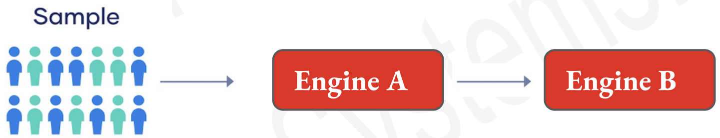
Few candidate approaches: each method must be tested on the same task.

1. Between Subjects (A-B testing)



- Easier to setup and analyse correctly
- No learning across conditions
- Test long term effects

2. Within Subjects

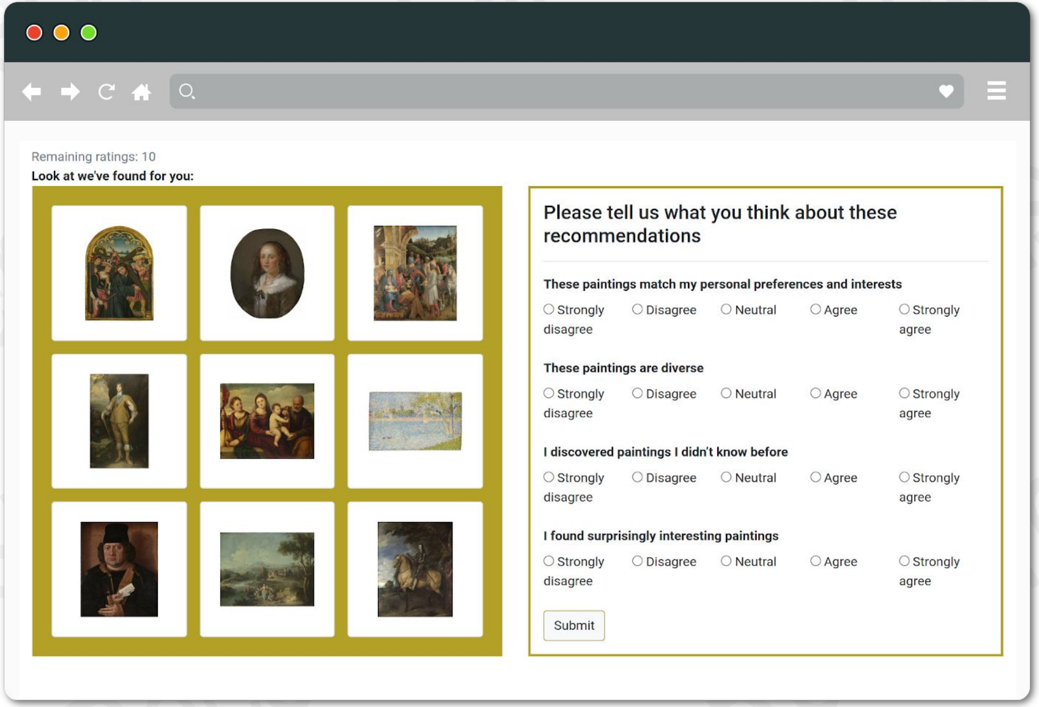


- More informative (superiority of methods)
- Can ask comparative questions about candidates

The RecSys pipeline: A case-study approach

Evaluation: User Study

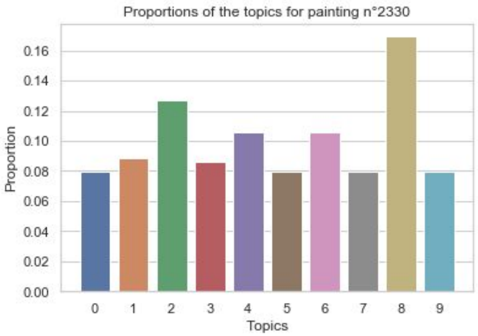
User-centric evaluation



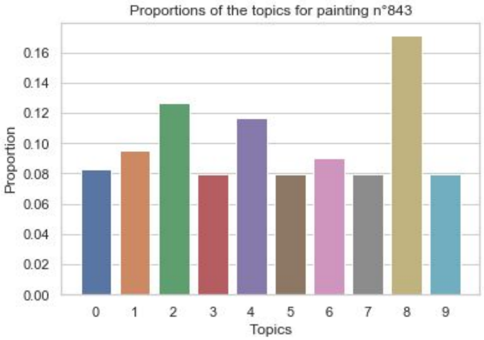
The RecSys pipeline: A case-study approach

Explaining Recommendations

Target painting



Most similar painting

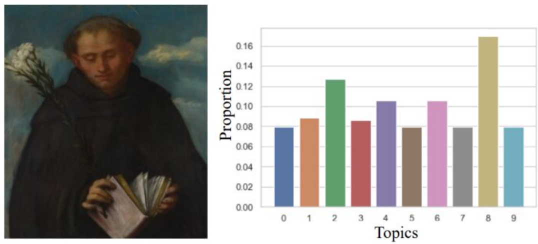
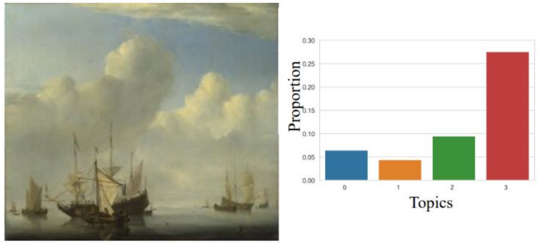
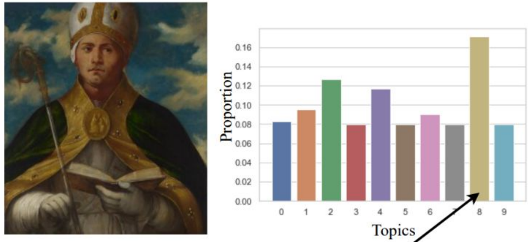
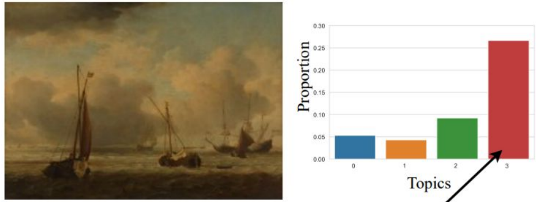


Topic 8
CHRIST
SAINT
JESUS
EVANGELIST
CROSS
CHURCH

Explainable recommendations have a positive impact on user experience.

The RecSys pipeline: A case-study approach

Explaining Recommendations

	LDA	BERT	ResNet
Target painting			
Most similar painting			

(christ, saint, altarpiece, panel, Jesus, new testament, evangelist, cross, church crucification)

(landscape, oil, van, anchor, 17th_century, river, view, scene, 17th_century landscape)

The RecSys pipeline: A case-study approach

UNIVERSITÉ DU LUXEMBOURG

Explaining Recommendations

Beauty

Old Age

Time

Batoni intends to encourage considering the brevity of youth and the inevitable passing of time.

Time orders Old Age to destroy Beauty
by Pompeo Girolamo Batoni
18th century

Semantic Gap

Mary Magdalene

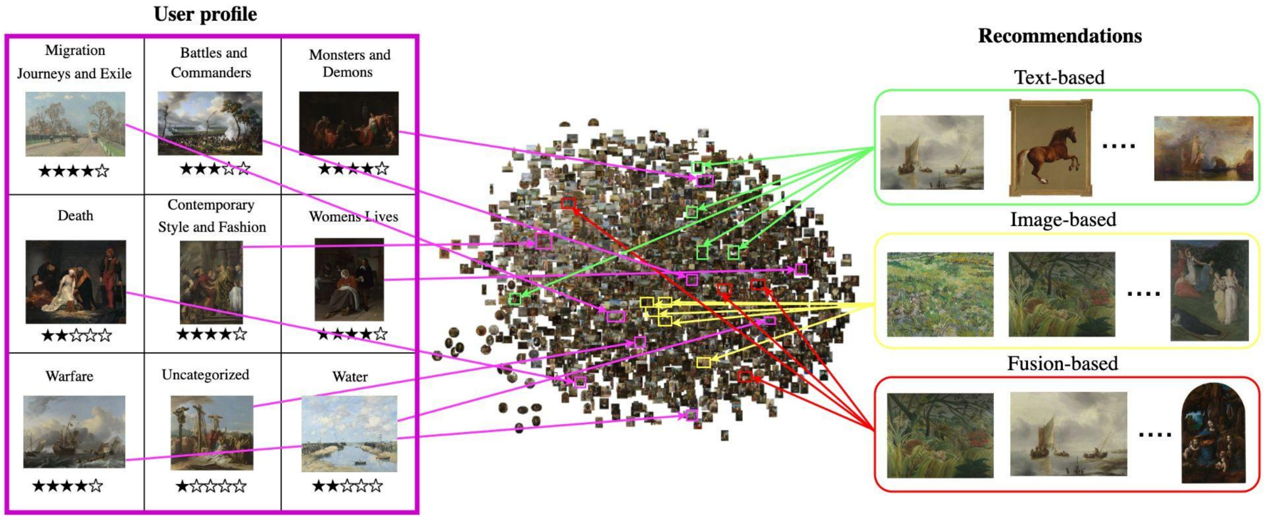
precious ointment

Religious

The Donor and Saint Mary Magdalene
by Marten van Heemskerck.
16th century

The typical RecSys Pipeline

Data Pre-processing → Model Training → Post Processing → Evaluation



Bereket A. Yilma and Luis A. Leiva: **The Elements of Visual Art Recommendation: Learning Latent Semantic Representations of Paintings for Personalized Recommendation,** Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI 2023)

Open Challenges

Formulating a RecSys Problem

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

Formulating a RecSys Problem

Data Pre-processing →  → Model Training → Post Processing → Evaluation

- Sort
- Filter
- Recommend

Understanding the context of the problem!

Multi-stakeholder Issues

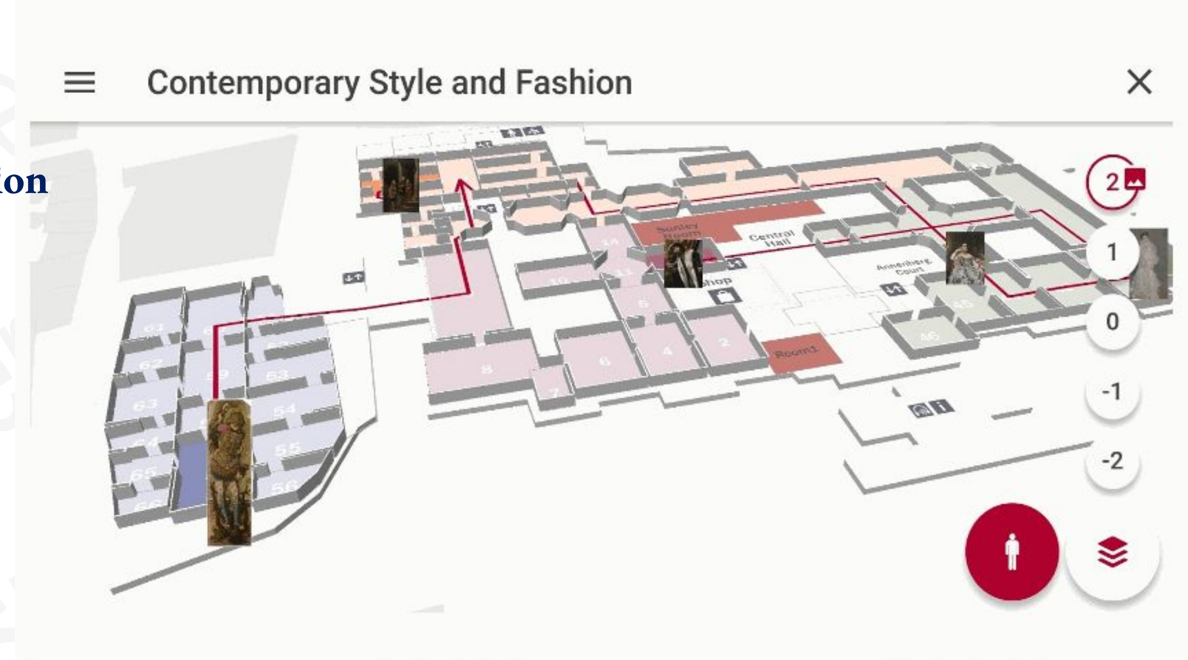
Why does understanding the context matter?

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

Multi-Stakeholder aware RecSys

1. POI (painting) Recommendation

2. Path Recommendation



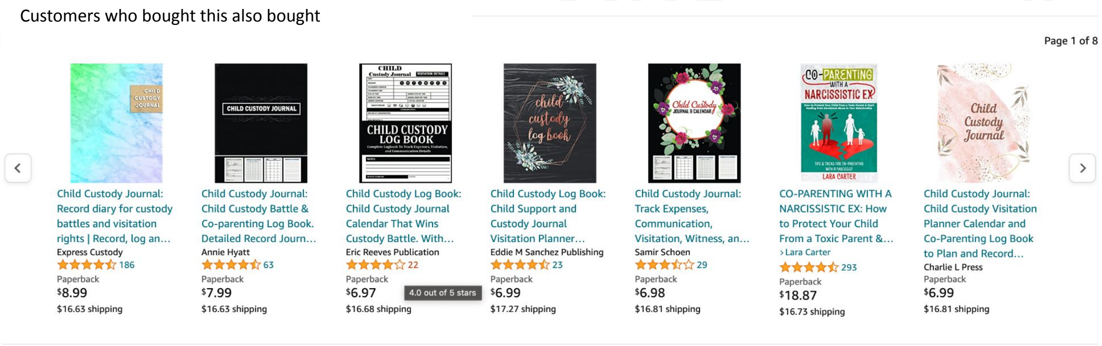
Multi-Stakeholder aware RecSys

Curator-visitor tradeoff

Issues and Challenges in RecSys

RecSys Issues & Challenges

- Adaptivity: changing business needs
- Robustness: Attack/ stress
- Privacy : Third party, sensitive private information

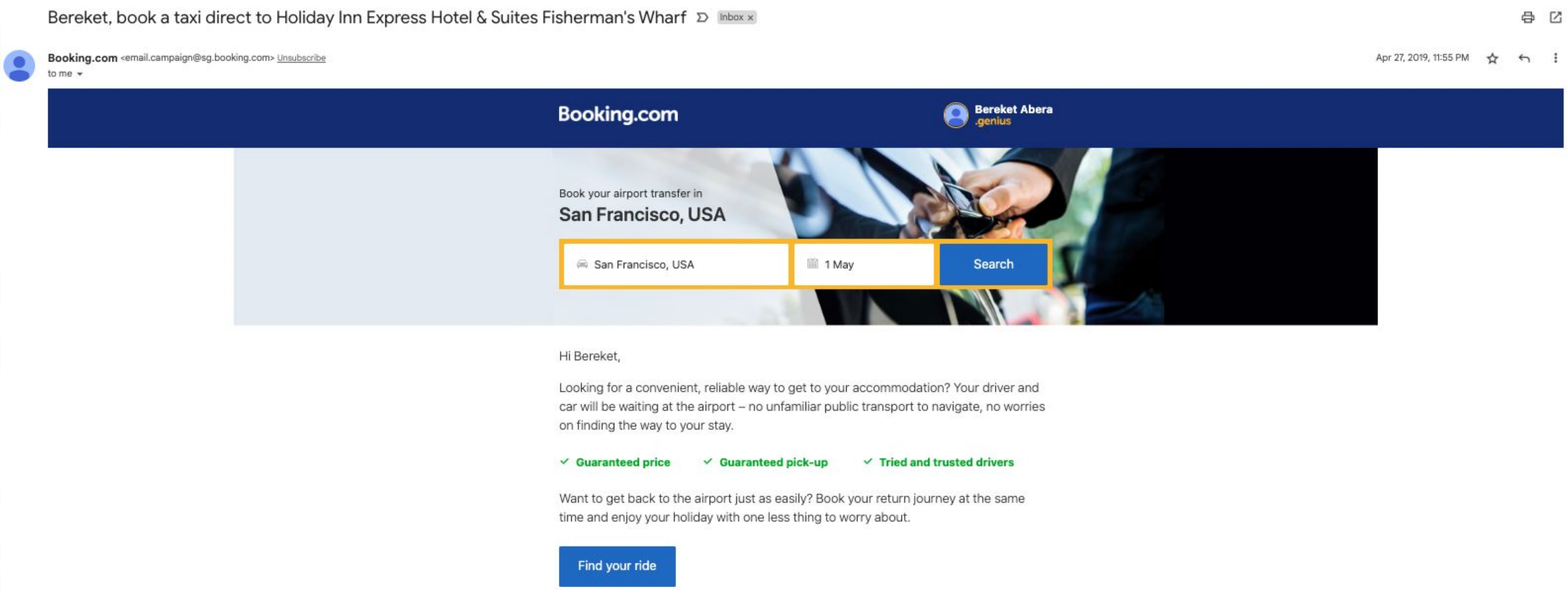


RecSys Issues & Challenges

RECOMMENDER SYSTEMS

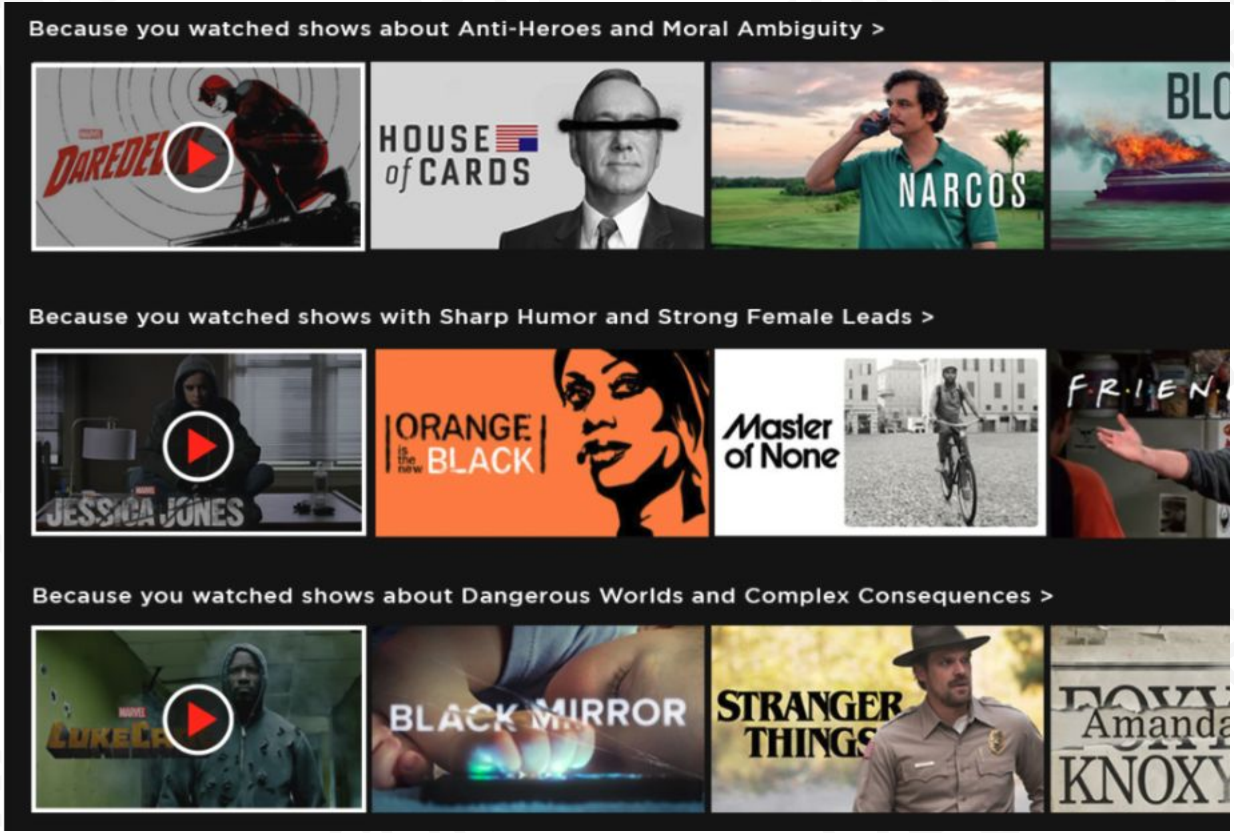
UNIVERSITÉ DU LUXEMBOURG

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

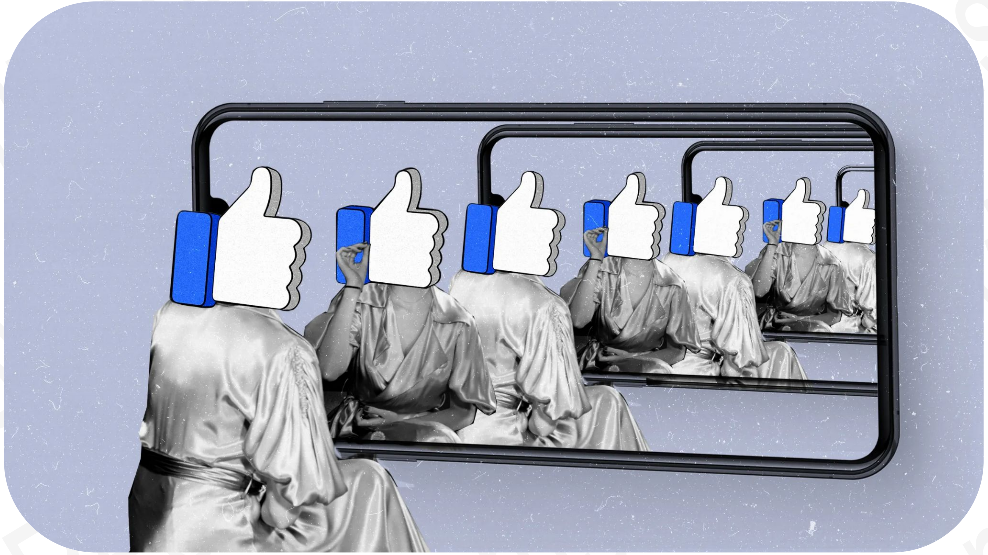


RecSys Issues & Challenges

- Explainability



Echo Chamber



Filter bubbles

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

RECOMMENDER SYSTEMS

UNIVERSITÉ DU LUXEMBOURG

“Recommender systems should not make our world smaller but bigger”

Thank You!

Contact

email: bereket.yilma@uni.lu

Website: https://bekyilma.github.io/

References

- **Bereket Abera Yilma,** Yannick Naudet and Hervé Panetto. "*Personalisation in Cyber-Physical-Social Systems: A Multi-stakeholder aware Recommendation and Guidance,*". In the proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization (UMAP '21), June 2021, Utrecht, Netherlands.
- **Bereket Abera Yilma,** Yannick Naudet and Hervé Panetto. "*Personalised visual art recommendation by learning latent semantic representations*". In the proceedings of t5th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP2020), October 2020Zakynthos Greece.
- **Bereket A. Yilma** and Luis A. Leiva: "*The Elements of Visual Art Recommendation: Learning Latent Semantic Representations of Paintings for Personalized Recommendation*", Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI 2023) April 2023, Hamburg ,Germany
  - Additional reading
    - Adversarial Attacks Against Visually Aware Fashion Outfit Recommender Systems
    - Privacy-Preserving Multi-View Matrix Factorization for Recommender Systems