NSWI166 – Introduction to Recommender Systems and User Preferences

Ladislav Peska & Peter Vojtas
Ladislav.peska@matfyz.cuni.cz, S208
https://www.ksi.mff.cuni.cz/~peska/vyuka/NSWI166

Lecture: Wed 9:00 S5 (:-/)
Labs: Wed 10:40 S6 (every two weeks)

2/1, ZK+Z, 4 credits

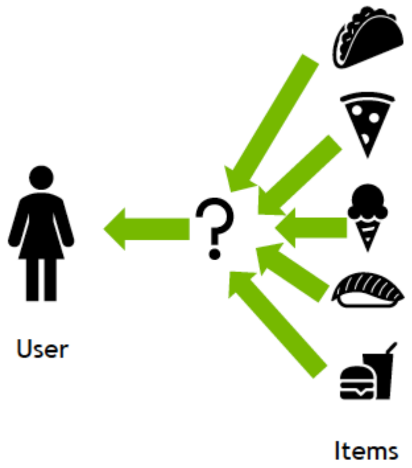
Today's Agenda

- **Introduction**
  - Problem domain
  - Purpose and success criteria
  - Paradigms of recommender systems
    - Collaborative Filtering
    - Content-based Filtering
    - Knowledge-Based Recommendations
    - Hybridization Strategies
- **NSWI166 outline**
  - Topics to be covered
  - Requirements
  - Dates
- **RecSys „Hello Worlds"**
  - Non-personalized
  - User-based KNN

Introduction

What are Recommender Systems?

*RS are **software agents** that **elicit** the **interests** and preferences of individual consumers […] **and make recommendations** accordingly. They have the potential to support and improve the quality of the decisions consumers make while searching for and selecting products online.* (Xiao & Benbasat 2007[1])



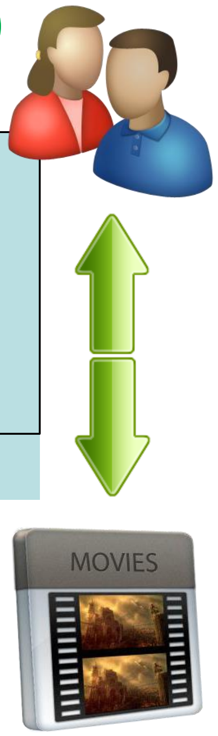
https://www.nvidia.com/en-us/glossary/recommendation-system/

[Wikipedia] A **recommender system** is a subclass of information filtering system that **provides suggestions** for items that are most **pertinent to a particular user**. Recommender systems are particularly useful when an individual needs to choose an item from a potentially overwhelming number of items that a service may offer.
[pertinent: relevant or applicable to a particular matter]

[N. Tintarev (KEN3160)] Recommender systems play an important role in **helping to mediate** many of our everyday **decisions and choices**….They do this by learning from our **past interactions**, inferring our interests and documenting our **preferences**. To make the right suggestions …recommender systems must …understand …**our current needs** and perhaps our **immediate intent**.

What are Recommender Systems?

- **Recommendation systems (RS) help to match users with items**
  - Ease information overload
  - ***Behave as a good sales assistent*** (guidance, advisory, persuasion,…)

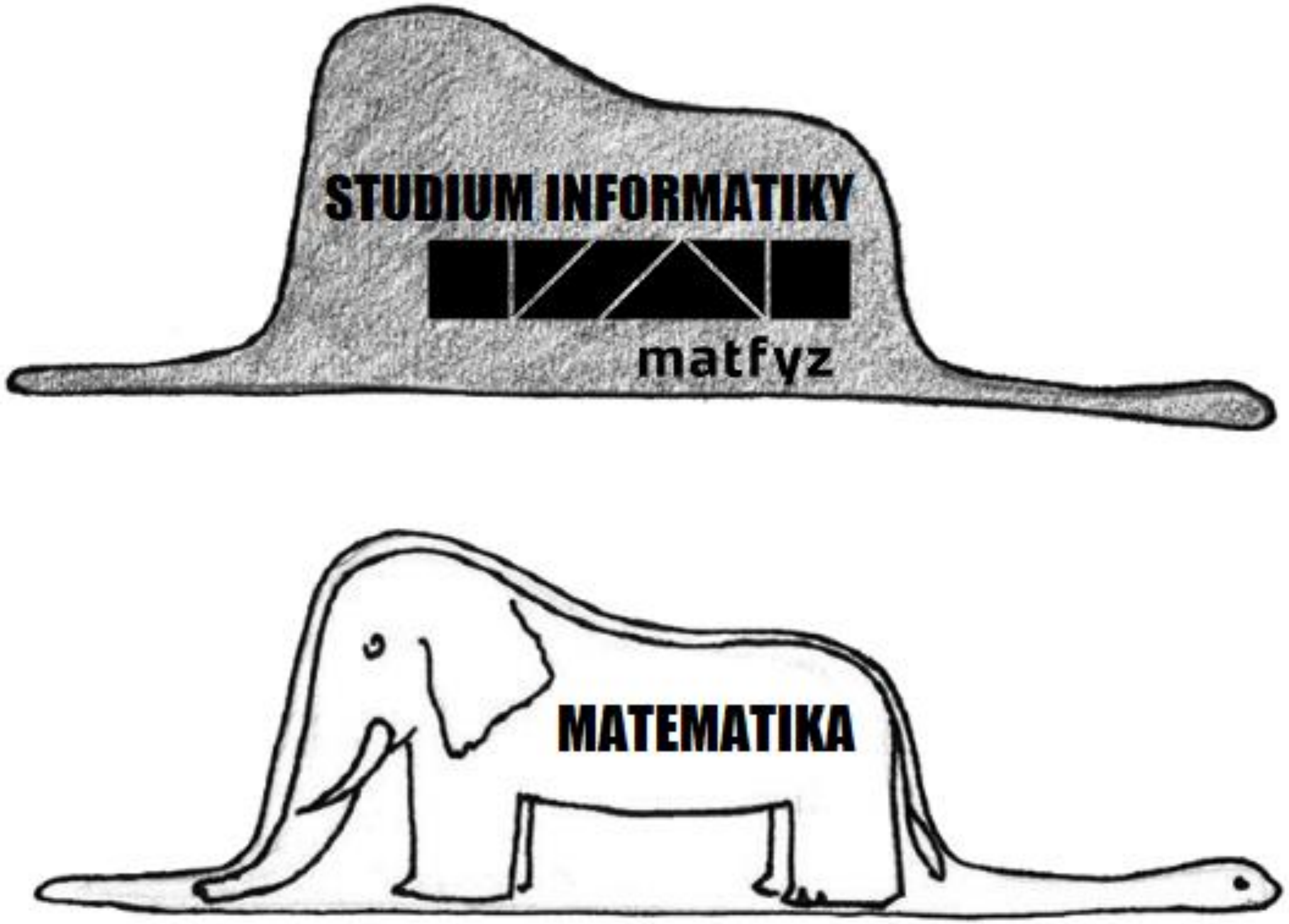


> *RS are software agents that elicit the interests and preferences of individual consumers […] and make recommendations accordingly.*
>
> *They have the potential to support and improve the quality of the decisions consumers make while searching for and selecting products online.*
>
> » (Xiao & Benbasat 2007[1])

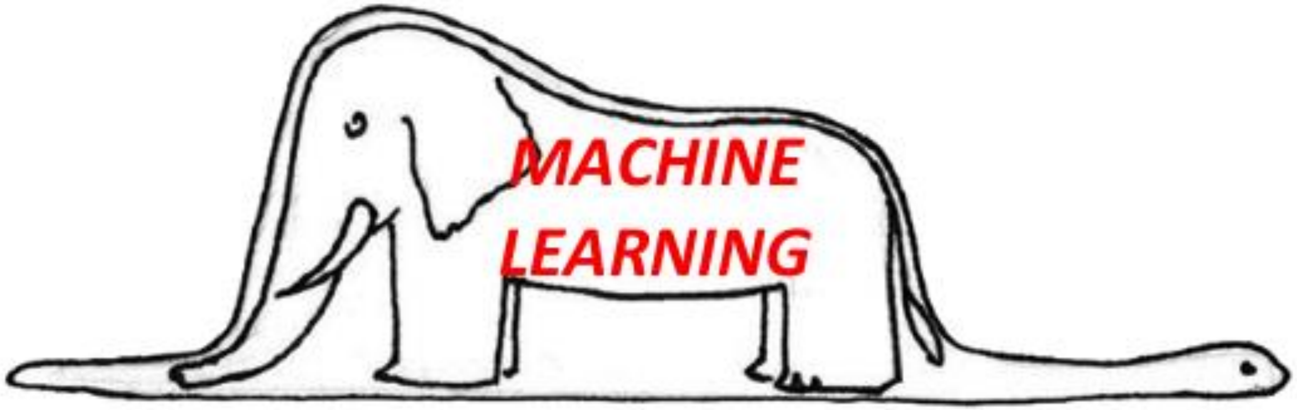
- **Different system designs / paradigms**
  - Based on availability of exploitable data
  - Implicit and explicit user feedback
  - Domain characteristics

(1) Xiao and Benbasat, *E-commerce product recommendation agents: Use, characteristics, and impact*, MIS Quarterly **31** (2007), no. 1, 137–209

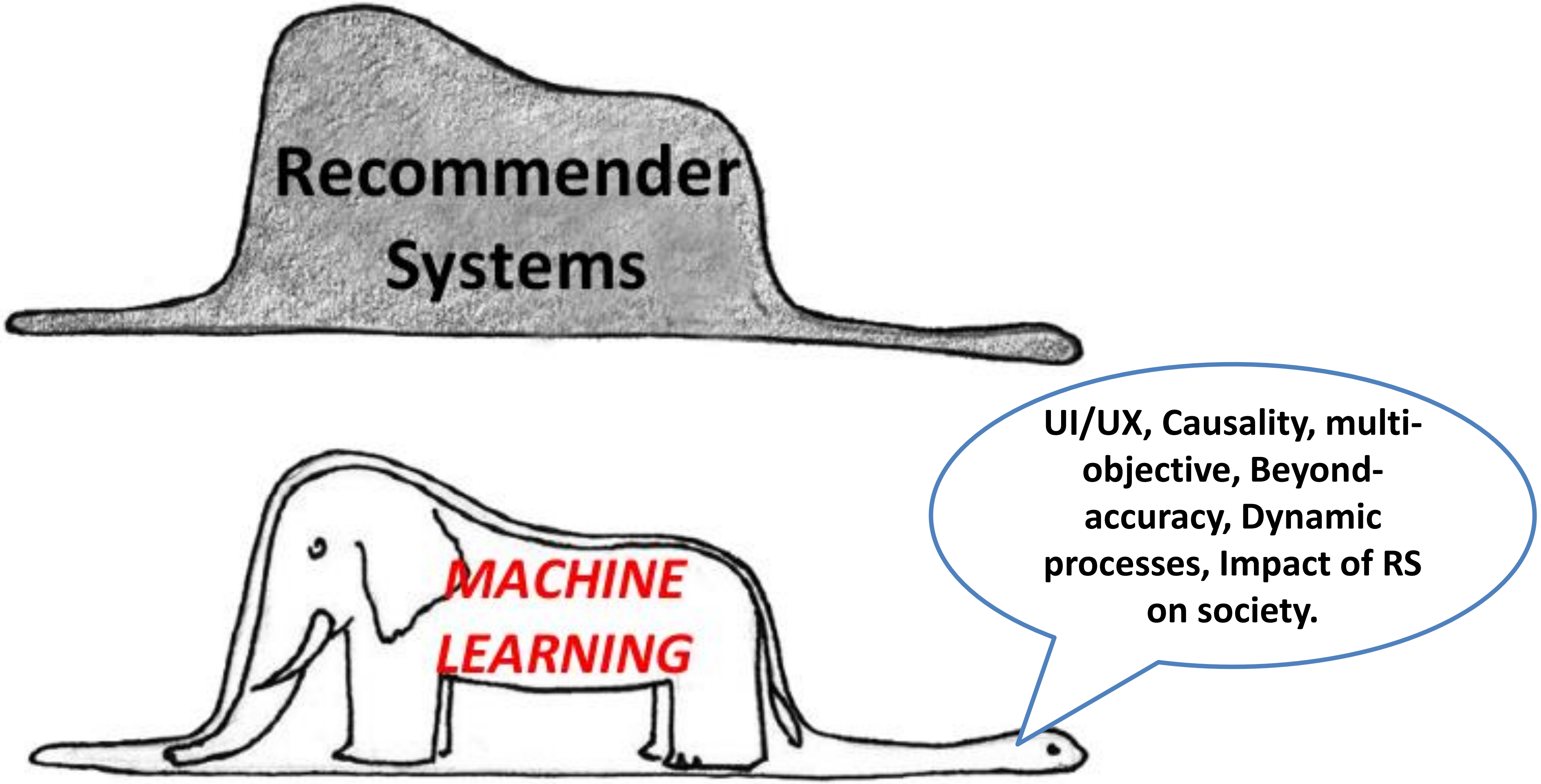
What are Recommender Systems? Meme style



What are Recommender Systems? Meme style



What are Recommender Systems? Meme style



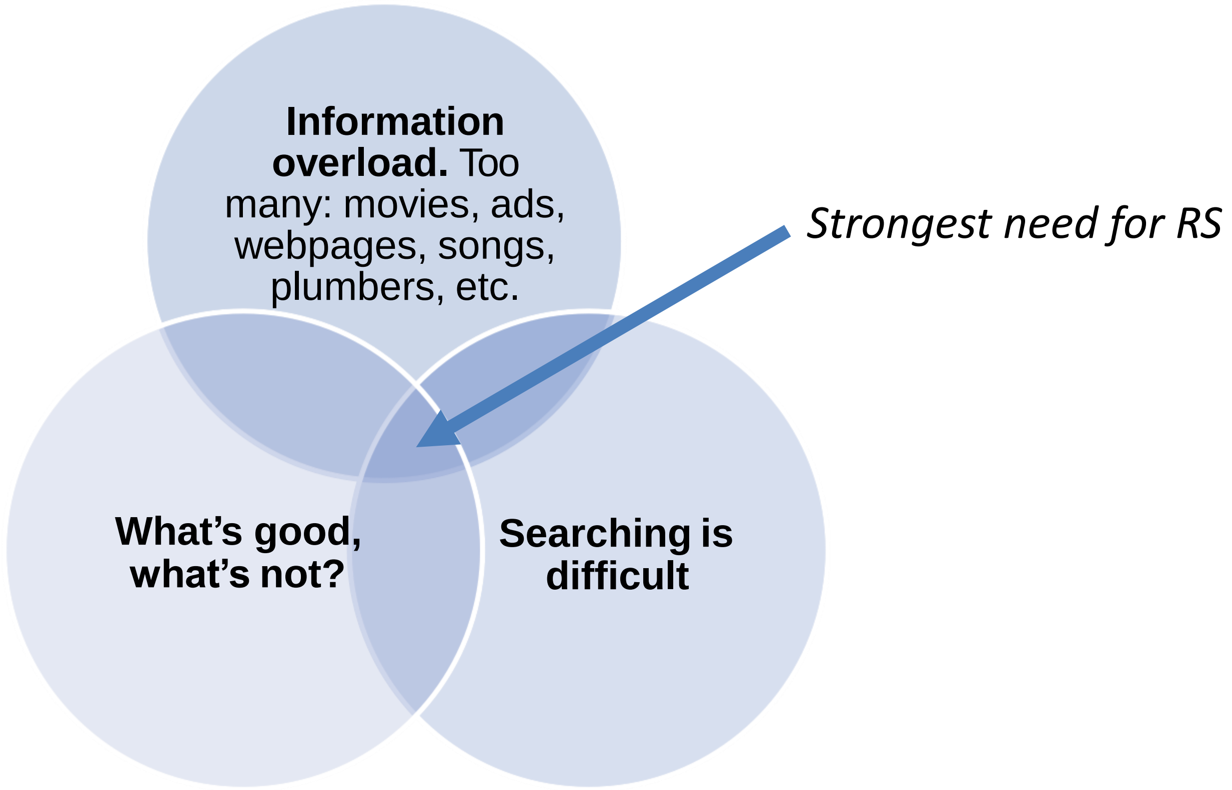
Recommending vs. Searching

- **Search Engines**
  - Users know in advance what they want *(and are able to specify it)*
  - ***Explicit query*** *submitted by the user*
  - *Evaluation through known „correct“ answers for the query*
- **Recommender Systems**
  - **Users do not know what they want** *(or do not know how to ask for it)*
  - ***RS tries to understand user's needs through observed behavior***
    (provide suitable results for these needs without being explicitly asked)
  - ***„Implicit query“***

Personalized Adds

- *Works extra-site, no prior indication of user's needs*
  - *Often, the basic principle of RecSys (mutual benefits) is violated*
  - *Although seemingly the same task, the same methods as for RecSys often do not work*

Benefits of recommendation



Benefits of recommendation

Information overload. Too many: movies, ads, webpages, son… plumbe…

…s

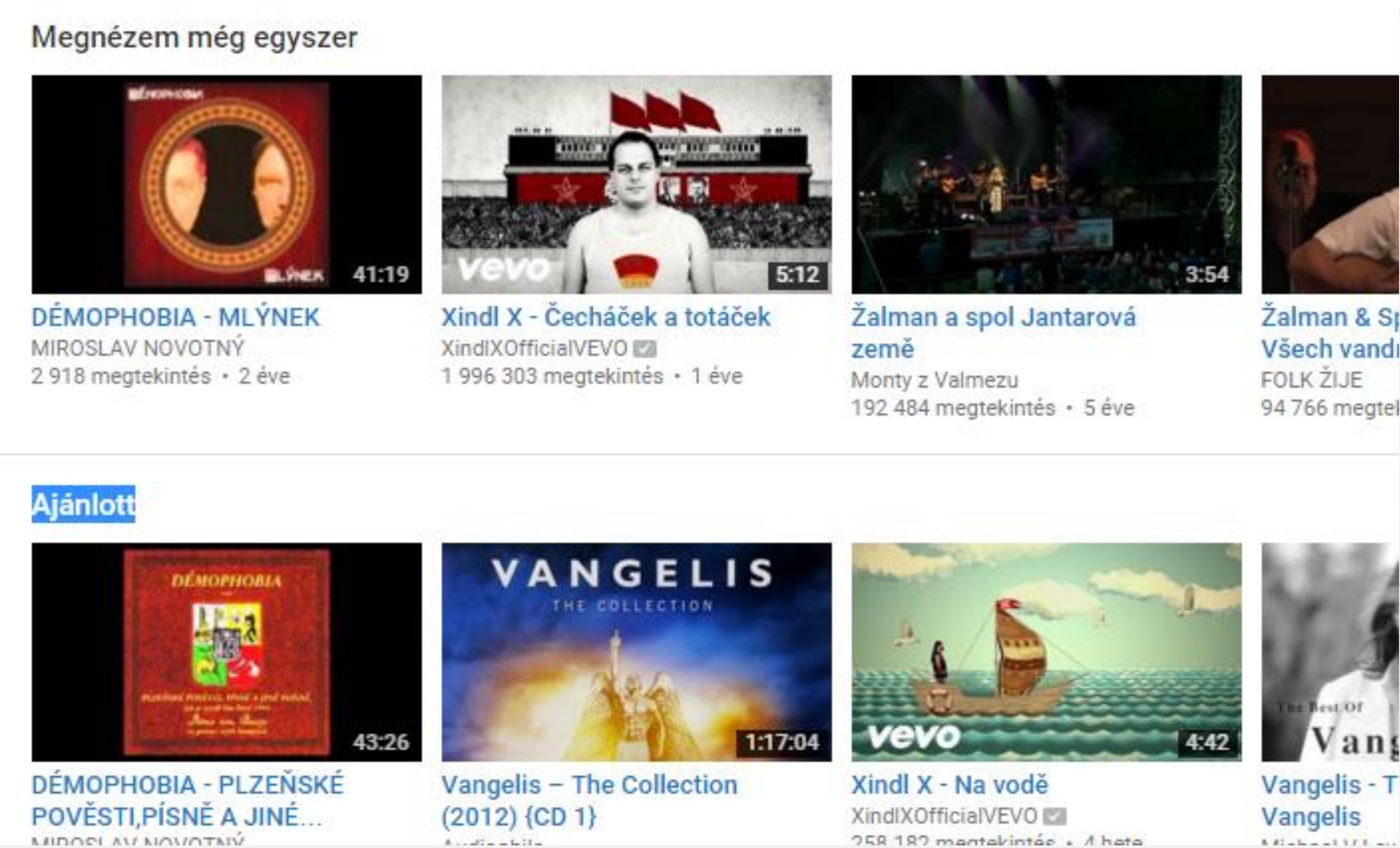
…d, …t's not?

Searching is difficult

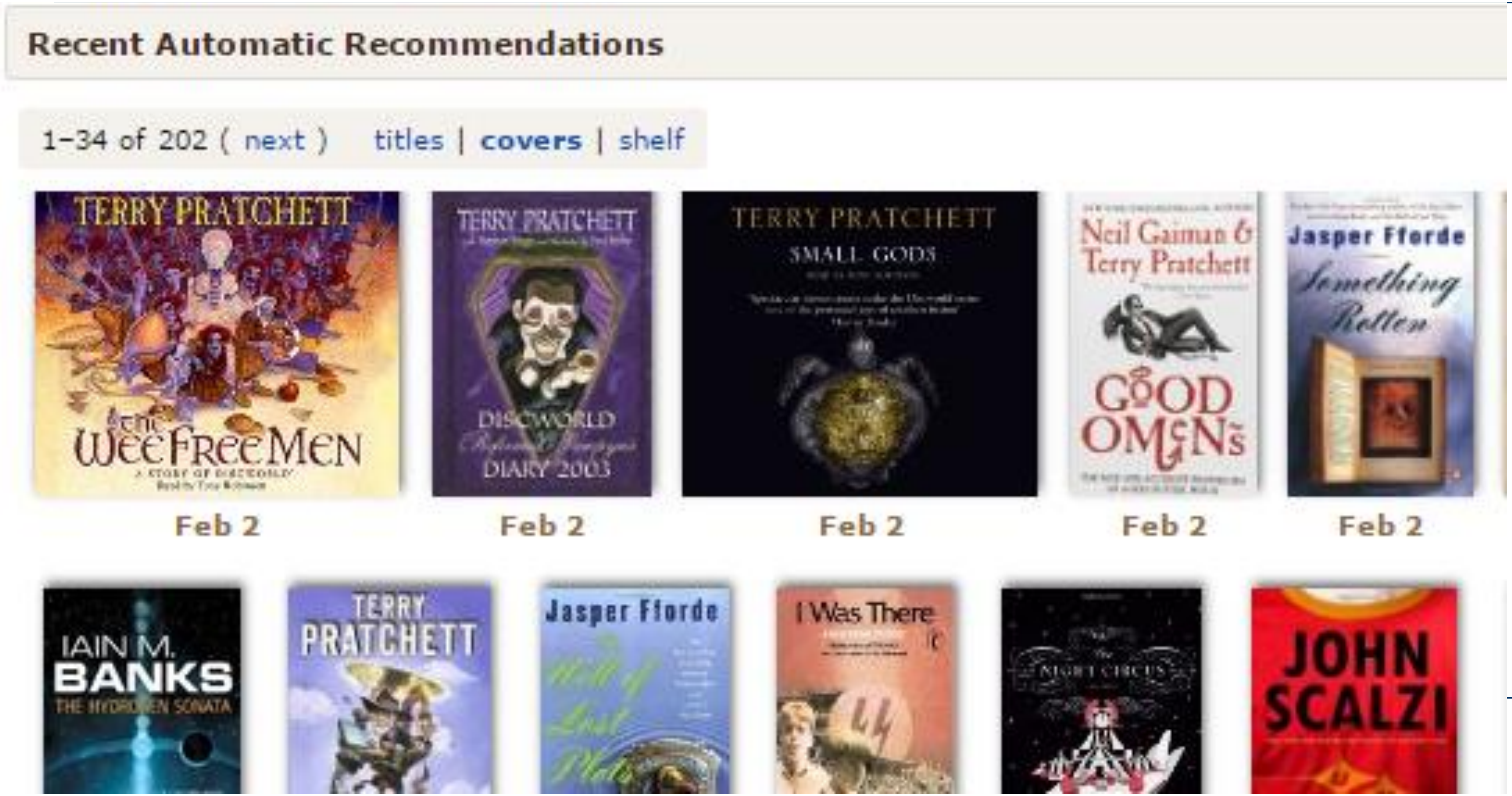
RecSys are everywhere

Movies


top picks
see more
based on your ratings, MovieLens recommends these movi
Band of Brothers
2001 R 705 min
Casablanca
1942 PG 102 min
One Fl
1975
CASABLANCA
recent releases
see more
movies released in last 90 days that you haven't rated
Cantinflas
Felony
What If
Frank
Sin City: A Dame to
If I Stay
Are


Megnézem még egyszer
41:19
DÉMOPHOBIA - MLÝNEK
MIROSLAV NOVOTNÝ
2 918 megtekintés • 2 éve
5:12
Xindl X - Čecháček a totáček
XindlXOfficialVEVO
1 996 303 megtekintés • 1 éve
3:54
Žalman a spol Jantarová země
Monty z Valmezu
192 484 megtekintés • 5 éve
Žalman & Sp
Všech vandr
FOLK ŽIJE
Ajánlott
DÉMOPHOBIA
43:26
DÉMOPHOBIA - PLZEŇSKÉ POVĚSTI,PÍSNĚ A JINÉ...
VANGELIS
THE COLLECTION
1:17:04
Vangelis – The Collection (2012) {CD 1}
vevo
4:42
Xindl X - Na vodě
XindlXOfficialVEVO
Vangelis - T
Vangelis

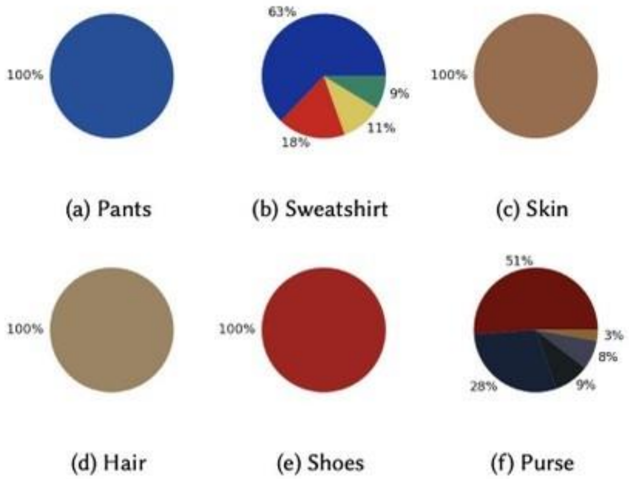
Books



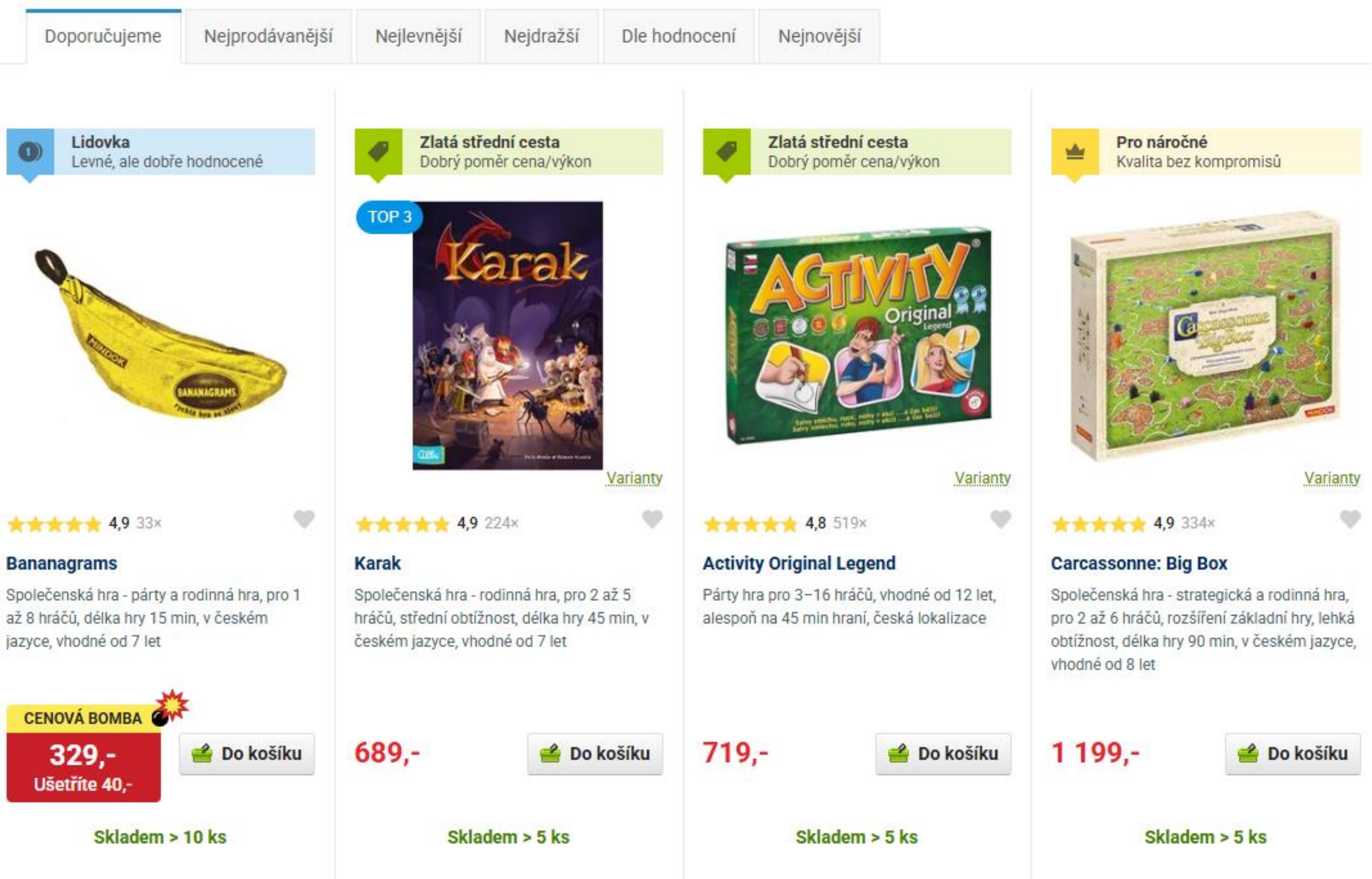
Recent Automatic Recommendations
1–34 of 202 (next) titles | covers | shelf
TERRY PRATCHETT
The Wee Free Men
TERRY PRATCHETT
DISCWORLD
DIARY 2003
TERRY PRATCHETT
SMALL GODS
Neil Gaiman & Terry Pratchett
GOOD OMENS
Jasper Fforde
Something Rotten
Feb 2
Feb 2
Feb 2
Feb 2
Feb 2
IAIN M. BANKS
THE HYDROGEN SONATA
TERRY PRATCHETT
Jasper Fforde
I Was There
JOHN SCALZI

Fashion recommendation

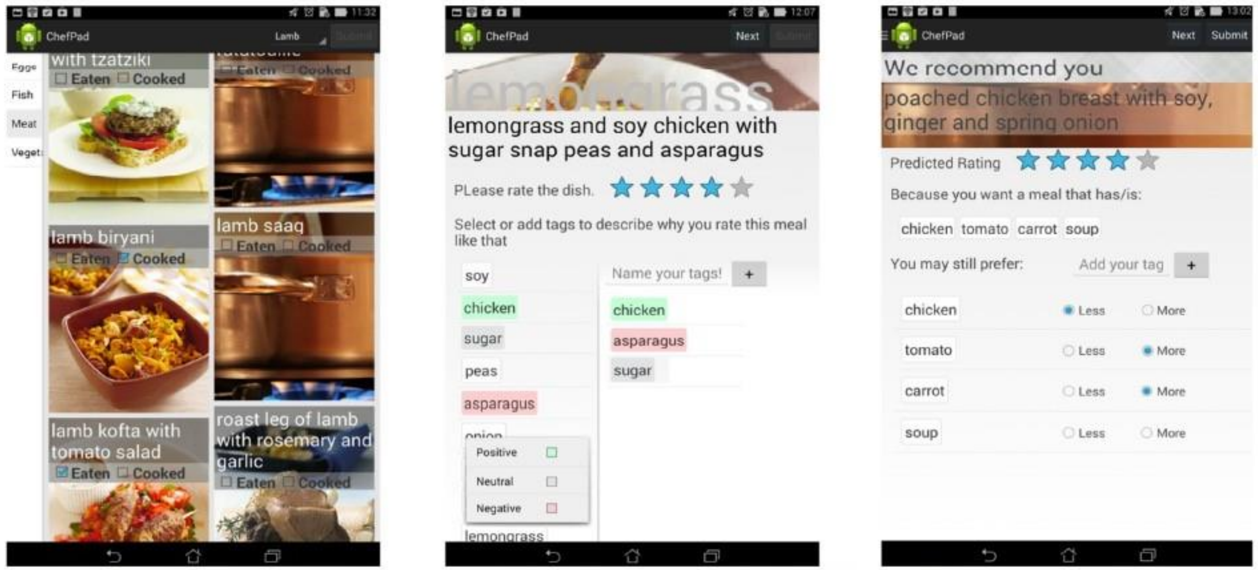
[Second Workshop on Recommender Systems for Fashion](#)



E-commerce in general



Food recommendations



Your Daily Mixes

Music

Play the music you love, without the effort. Packed with your favorites and new discoveries.

Daily Mix 1

Bachan Kaur, Edo & Jo, Ty Burhoe and more

MADE FOR NAVA

Daily Mix 2

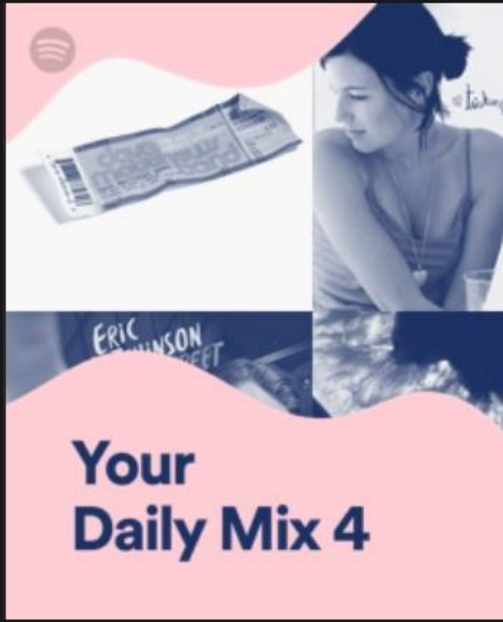
James Vincent McMorrow, Bon Iver, Angus & Julia Stone and more

MADE FOR NAVA

Daily Mix 3

Mika Mendes, C4 Pedro, DJ Dias Rodrigues and more

MADE FOR NAVA

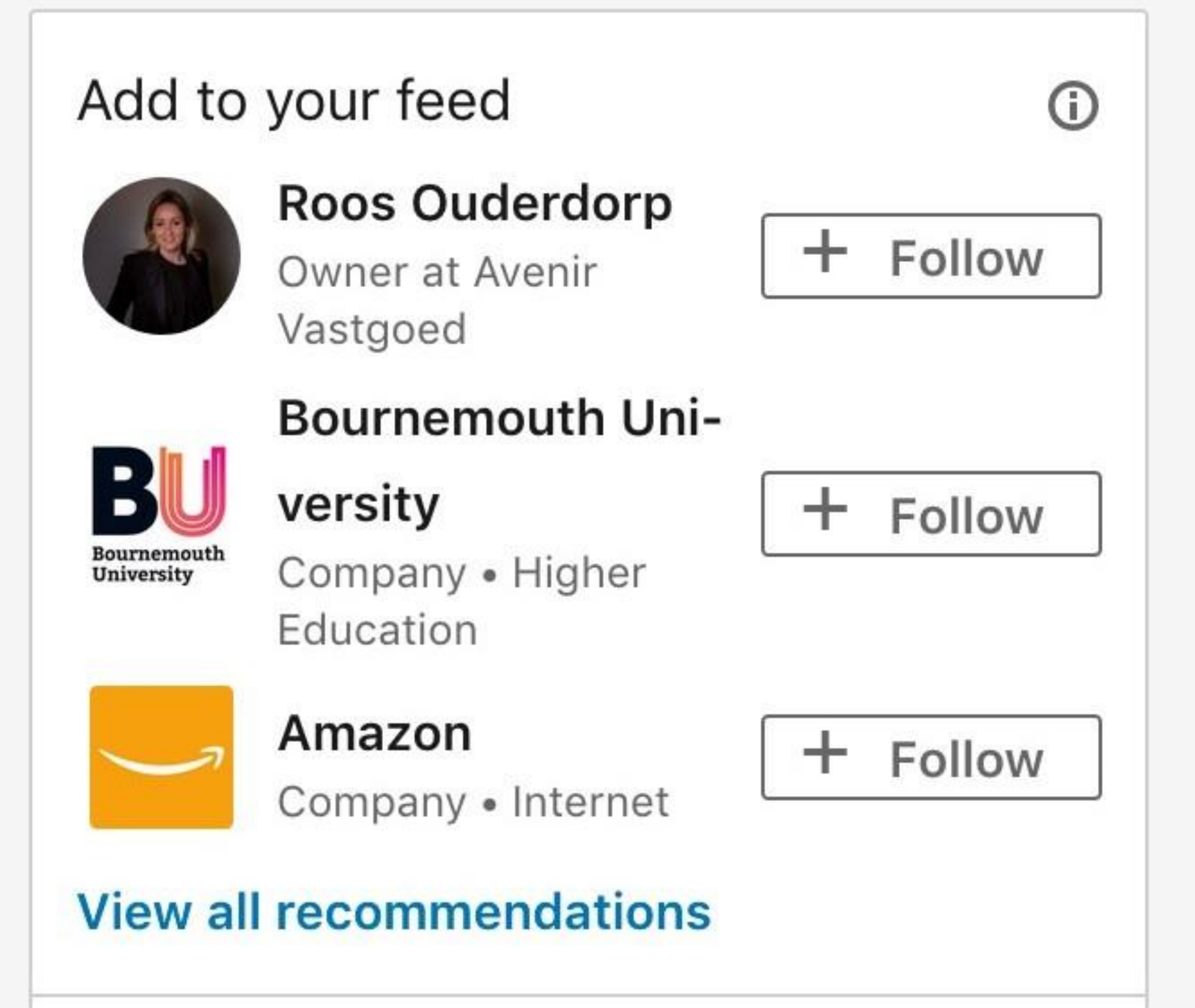


Daily Mix 4

Dave Matthews Band, Trista Prettyman, Eric Hutchinson more

MADE FOR NAVA

People (LinkedIn)



People (Bumble)

- Jobs
- News
- Contacts
- Points of interests
- …

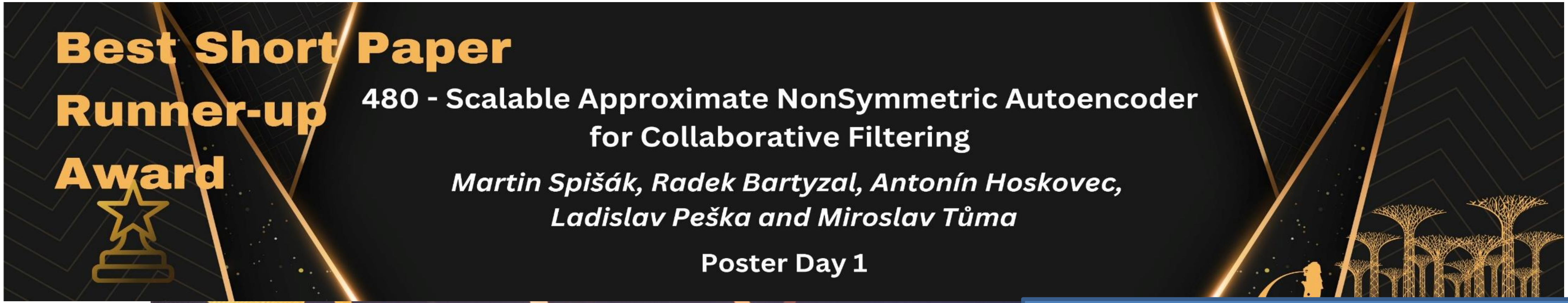
Recommender Systems are intensively researched

RecSys 2021: >1100 attendies, ±50 papers (acceptance ~20%)
Similar for **RecSys 2022-2024**

- Large industrial sponsors, lots of industry based research papers
- Lots of „We're hiring" signs ☺

LIST OF WORKSHOPS

- CARS: Workshop on Context-Aware Recommender Systems
- ComplexRec: Workshop on Recommendation in Complex Environments
- FAccTRec: Workshop on Responsible Recommendation
- FashionxRecSys: Workshop on Recommender Systems in Fashion and Retail
- GReS: Workshop on Graph Neural Networks for Recommendation and Search
- INRA: Workshop on News Recommendation and Analytics
- IntRS: Joint Workshop on Interfaces and Human Decision Making for Recommender Systems
- KaRS: Workshop on Knowledge-aware and Conversational Recommender Systems
- MORS: Workshop on Multi-Objective Recommender Systems
- OHARS: Workshop on Online Misinformation- and Harm-Aware Recommender Systems
- ORSUM: Workshop on Online Recommender Systems and User Modeling
- PERSPECTIVES: Workshop on Perspectives on the Evaluation of Recommender Systems
- PodRecs: Workshop on Podcast Recommendations
- RecSys Challenge Workshop
- RecSys in HR: Workshop on Recommender Systems for Human Resources
- RecTour: Workshop on Recommenders in Tourism
- SimuRec: Workshop on Synthetic Data and Simulation Methods for Recommender Systems Research
- XMRec: Workshop on Cross-Market Recommendation



This can be you ☺

RecSys 2023, Martin Spisak (NSWI166 almuni)

Purpose and success criteria (1)

Retrieval perspective

- Reduce search costs
- Provide "correct" proposals
- Users know in advance what they want

Recommendation perspective

- Serendipity – identify items from the Long Tail
- Users did not know about existence, but like them

Purpose and success criteria (2)

Prediction perspective

- Predict to what degree users like an item
- Used to be the most popular evaluation scenario

Interaction perspective

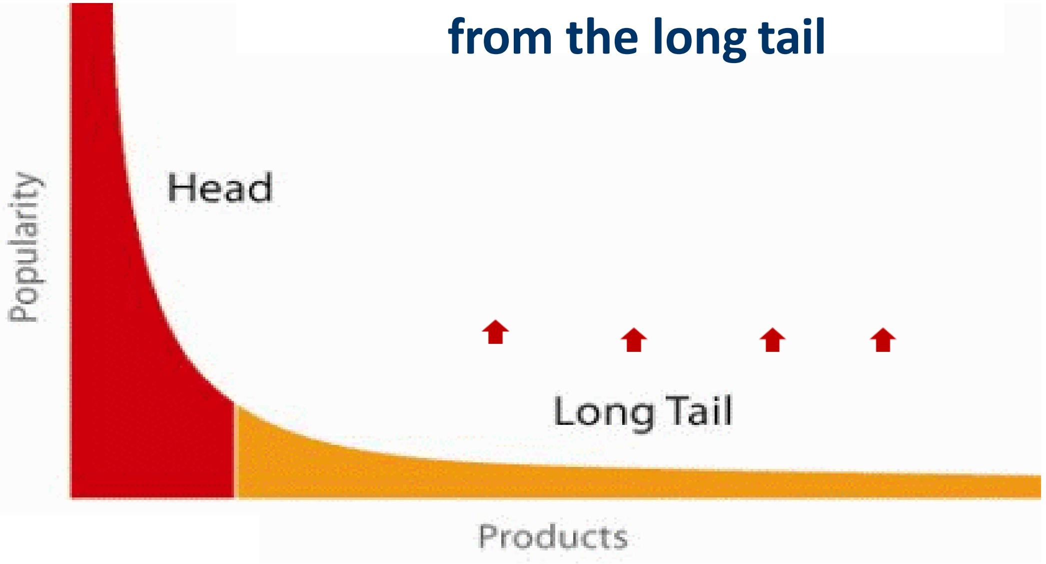
- Give users a "good feeling"
- Educate users about the product domain
- Convince/persuade users - explain

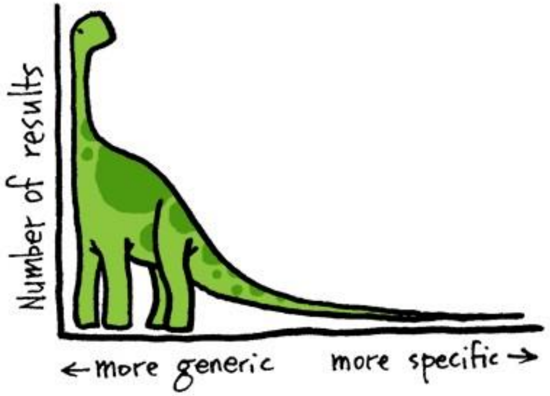
Finally, conversion perspective

- Commercial situations
- Increase "hit", "clickthrough", "lookers to bookers" rates
- Optimize sales margins and profit

When does a RS do its job well?

Recommend items from the long tail





- "Recommend widely unknown items that users might actually like!"
- 20% of items accumulate 74% of all positive ratings

Recommender systems

- **RS seen as a function**
- **Given:**
  - User model (e.g. ratings, preferences, demographics, situational context)
  - Items (with or without description of item characteristics, relations,…)
- **Find:**
  - Items’ relevance score / **Ranking** of items *(short head section is sufficient)*.

- **How?**
  - Based on **similarity**! *(and other stuff… later)*
  - *The true question is, what kind of similarity we assume*☺

Discuss

How do you think a recommender system works?
What could be similar?

Paradigms of recommender systems

Recommender systems reduce information overload by estimating relevance

Recommendation component

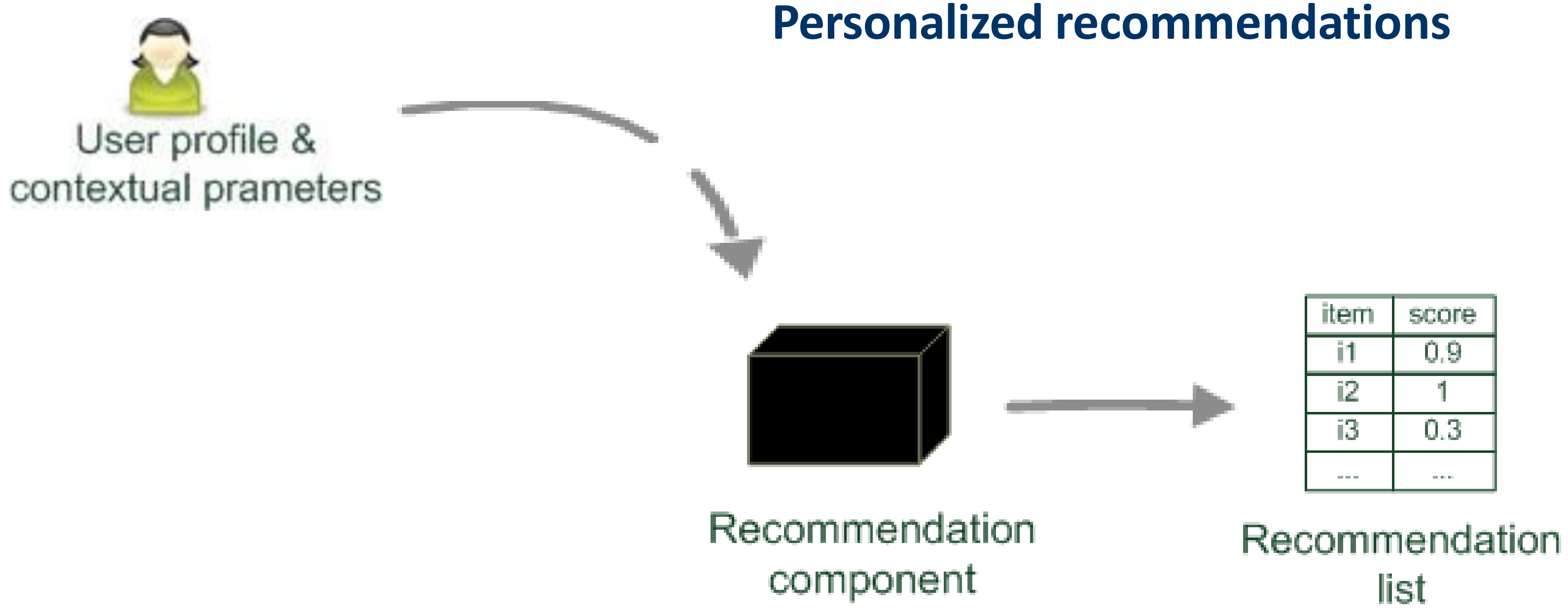
item	score
i1	0.9
i2	1
i3	0.3
...	...

Recommendation list

Non-personalized, e.g. most popular, manually currated

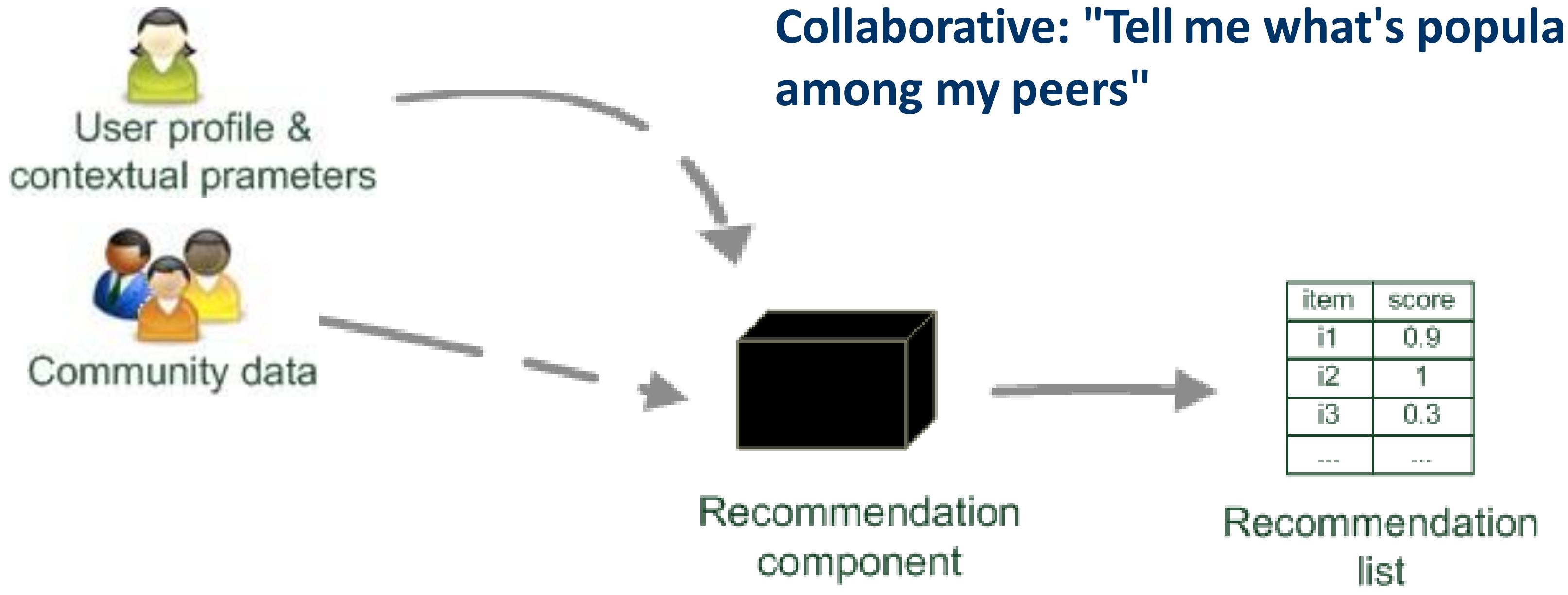
Paradigms of recommender systems

Personalized recommendations



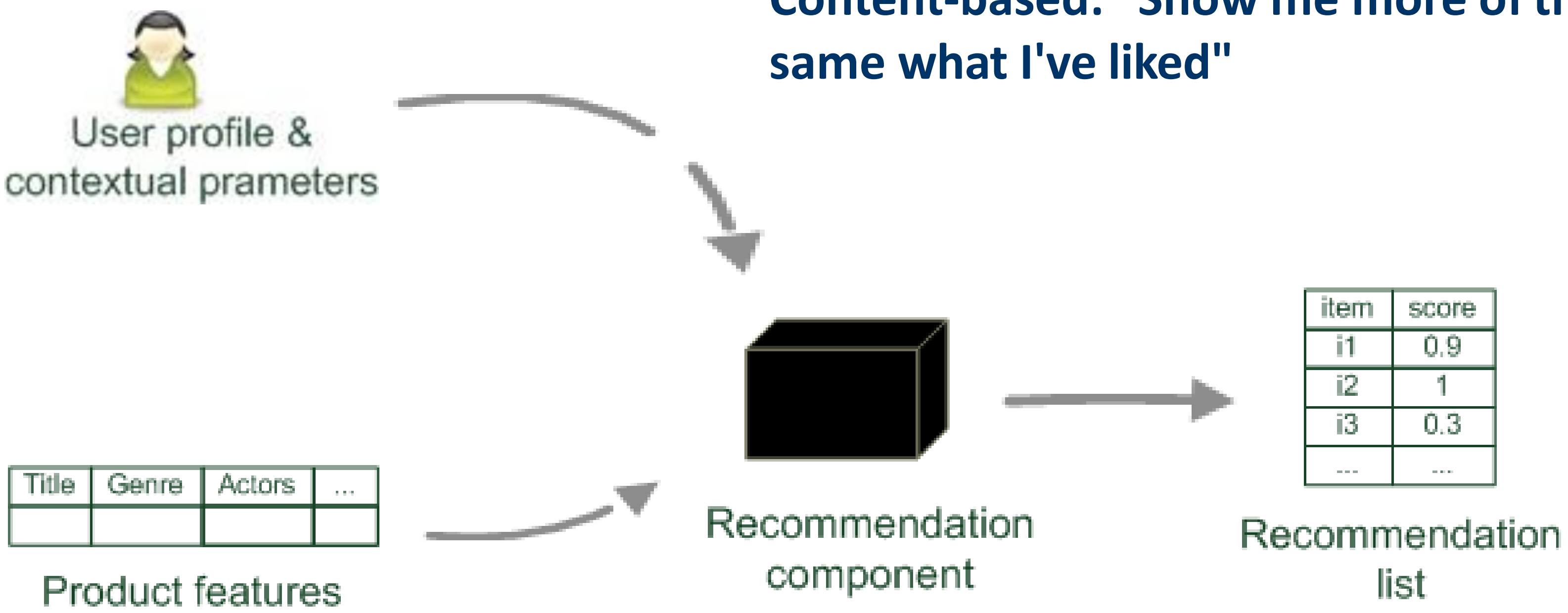
Paradigms of recommender systems

Collaborative: "Tell me what's popular among my peers"



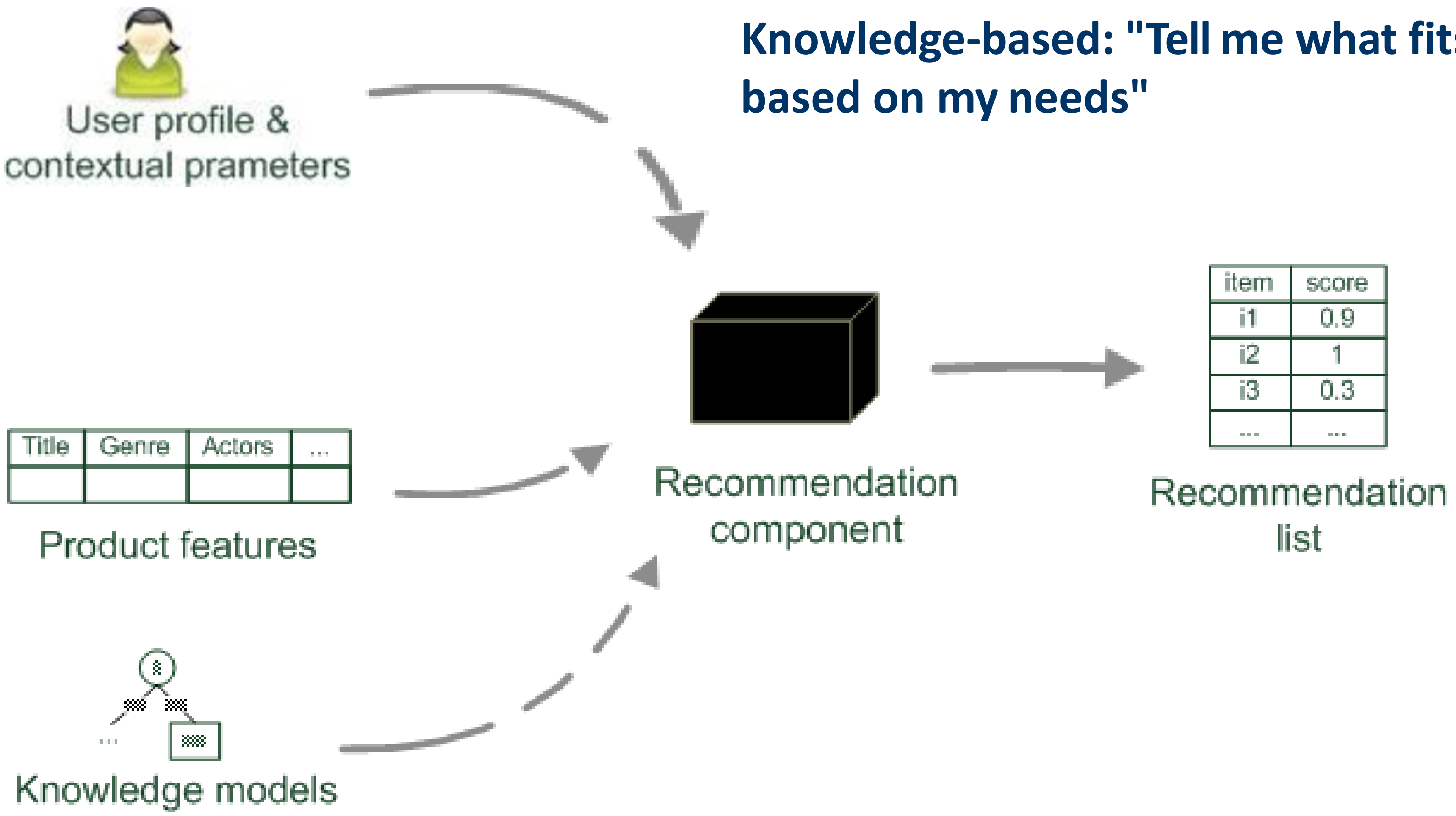
Paradigms of recommender systems

Content-based: "Show me more of the same what I've liked"



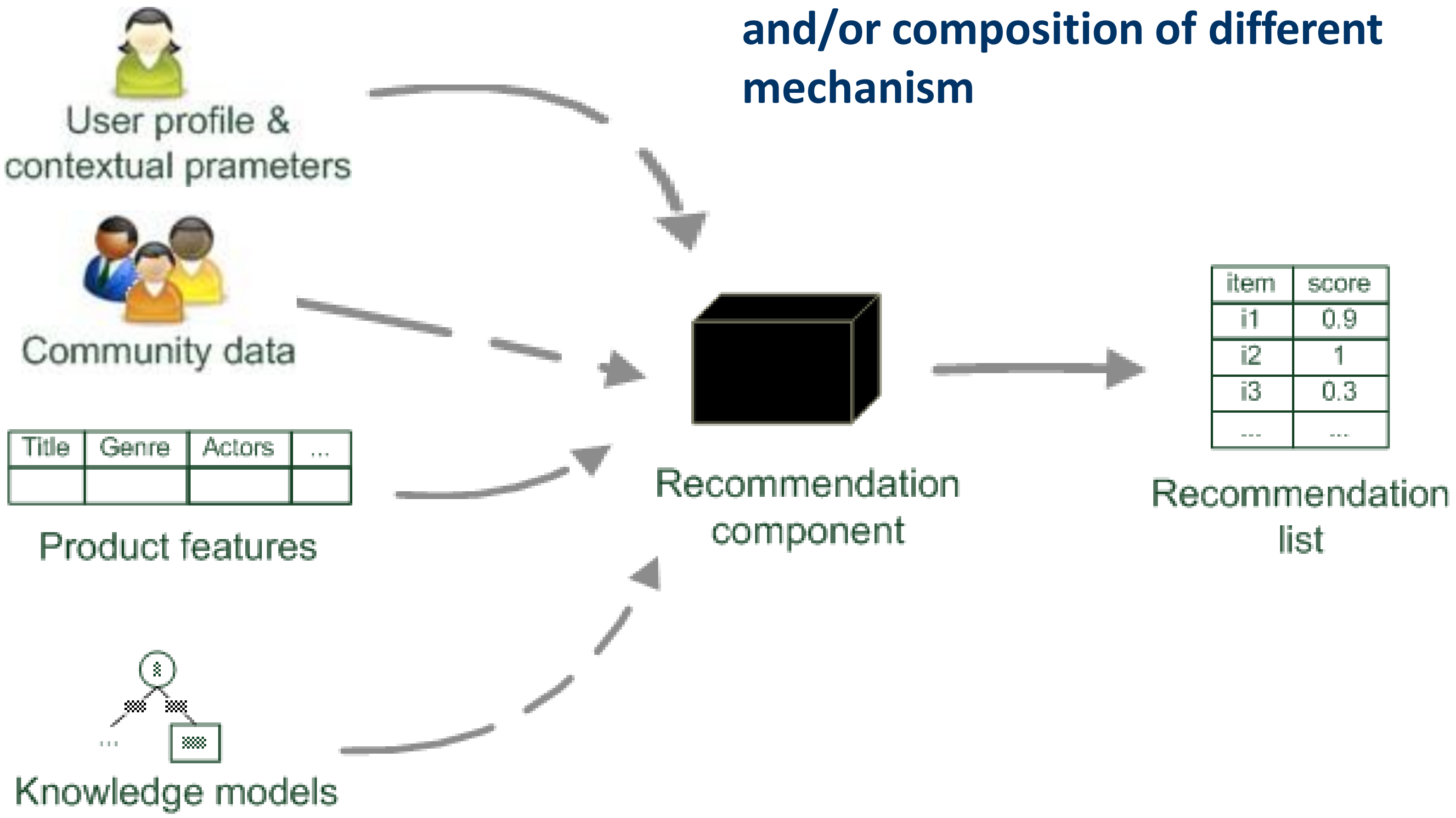
Paradigms of recommender systems

Knowledge-based: "Tell me what fits based on my needs"



Paradigms of recommender systems

Hybrid: combinations of various inputs and/or composition of different mechanism

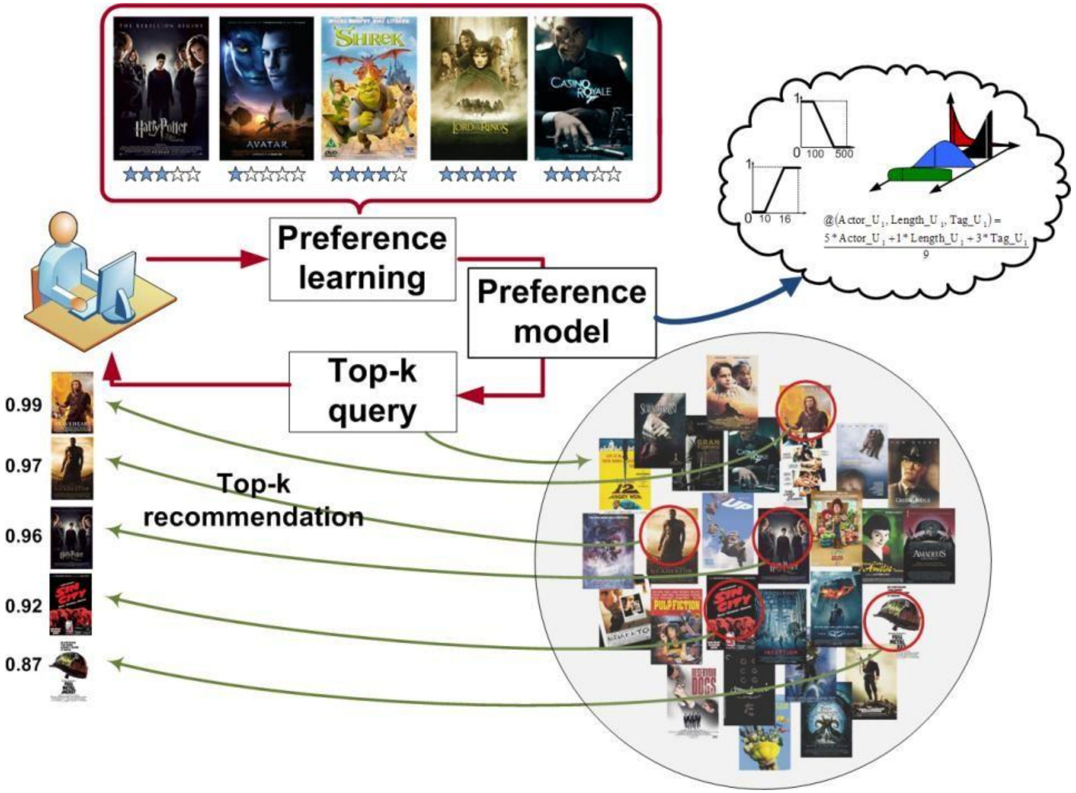


Lifecycle of Recommender Systems

1. Get User Feedback
2. Learn internal model
3. Upon demand, recommend objects

- The process is asynchronous by nature
- Most recent usually most relevant
- Dynamic nature of the process seriously complicate things

- Recommender systems may affect its users



„If you gaze long into an abyss, sometimes the abyss gazes back.“

Questions?

NSWI166 Lectures Outline

RS Basics

- **20.2.** Intro + RecSys Hello World
- **27.2.** Collaborative Filtering
- **06.3.** Collaborative Filtering (Intro to Content-based RS)
- **13.3.** Content-based and Knowledge-based RS
- **20.3.** Evaluating RS
- **27.3.** Evaluating RS

User preferences

- **03.4.** Explicit and Implicit Feedback
- **10.4.** Visualize user preferences

(Slightly more) advanced RS

- **17.4.** Hybrid recommending algorithms
- **24.4.** Hybrid recommending algorithms & beyond-relevance RS
- **15.5.** Deep Learning in RS and recent trends
- **22.5.** Invited Lecture (to be confirmed, Seznam.cz)

NSWI166 Labs

Lab Content:

- Get practical understanding of user preferences and recommendation concepts
- Development & evaluation of rec. algorithms
- Hands on Libraries / Tools / Frameworks / Datasets
- Time for paper/software presentations etc.

Active reading

- Four research papers
- Short written reports (strict deadlines)
- *At least two (accepted) reports to pass; additional/exceptional reports = bonifications for exams*

Participation during labs (or at home)

- Typically, 2-3 assignments will require to be finalized at home
- *At least 50% of assignments accepted to pass; additional/exceptional results = bonification for exams*

Semestral work (individual or in a group)

- Details after the first block
- User study to compare recommending algorithms
  - Integration of existing recommending frameworks
  - Slight results modification (e.g., diversity enhancments)
  - Integration to EasyStudy
  - Non-trivial domain / integration / algorithms => group work

Alternative pass: more extensive research project (contact me for details)

Active reading

#1: Collaborative RS algorithms (choose one of the following)

- **Amazon.com Recommendations Item-to-Item Collaborative Filtering:**
  - https://www.cs.umd.edu/~samir/498/Amazon-Recommendations.pdf

- **Embarrassingly Shallow Autoencoders for Sparse Data**
  - https://dl.acm.org/doi/abs/10.1145/3308558.3313710

- **Collaborative filtering with temporal dynamics**
  - https://dl.acm.org/doi/abs/10.1145/1557019.1557072

Deadline: 10.3.2024

At most 1 page of text CZ/SK/EN (normal font and margins - be brief but thorough)

Q1. **Your name and Paper's title**

Q2. **What is this paper about, and what contributions does it make?**

Q3. **What main new insights YOU received from the paper?**

Q4. **Does the paper has any notable weaknesses/limitations?**

Exam

„Semi-oral“ exam (written notes + discussion)

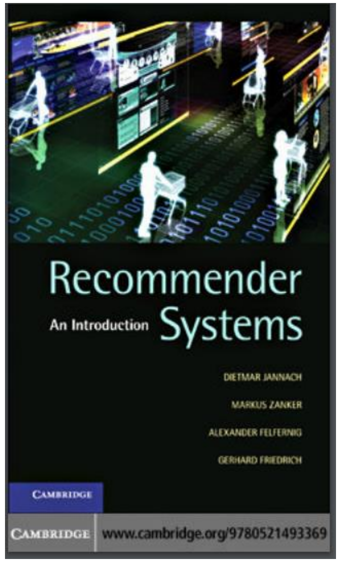
- 4-5 questions from the theory covered during lectures
- Application of knowledge
    - *Having a particular case where RS is to be introduced, you should be able to recommend sensible first approaches & hypotheses on what could work, what not.*

    - *Sample questions later during the course.*

Source materials

Recommender Systems
– An Introduction

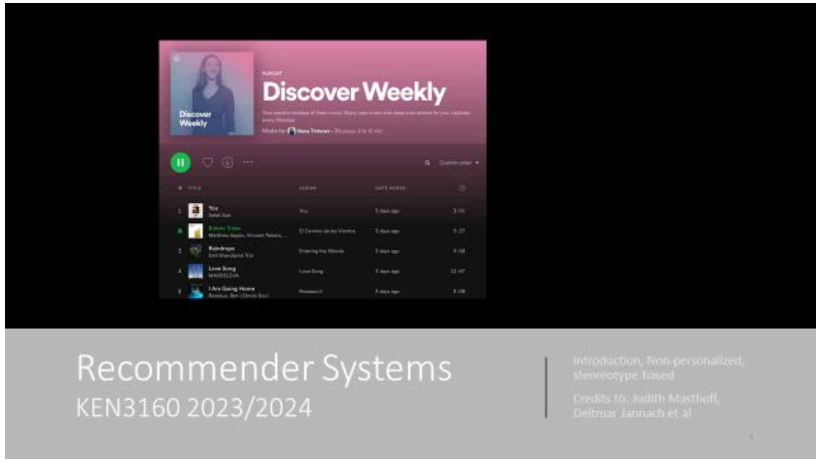
Dietmar Jannach, Markus Zanker,
Alexander Felfernig, Gerhard Friedrich
Cambridge University Press
- Bit outdated, but covers the basics



KEN3160: Recommender Systems
Maastricht University

Nava Tintarev, Francesco Barile

- Borrowed some content



Ad-hoc papers
&tutorials

NSWI166 - Literature

- **Textbooks**
  - Ricci, F. et al (Eds): Recommender Systems Handbook, Springer, 2011/2015/2022
    - https://edyaaleh.files.wordpress.com/2016/02/recommendersystemshandbook.pdf
    - Mail me for the most up-to-date book's PDF
  - Kim Falk: Practical Recommender Systems, Manning 2019
- **Tutorials**
  - https://www.slideshare.net/balazshidasi (algorithms, deep learning)
  - https://www.slideshare.net/usabart (evaluation)
  - RecSys 2021+2022+2023 tutorials (let me know if you cannot obtain videos for tutorials)
  - NSWI166 slides: http://ksi.mff.cuni.cz/~peska/vyuka/nswi166/
- **Lecture recordings from last years (in Czech)**
  - https://www.youtube.com/playlist?list=PLTYzSYF3EX3AO5IouvxraVzeVPRORJTmV
- **Conferences / Journals / Other sources**
  - **RecSys**, **UMAP**, **SIGIR**, IUI, CHI, ECIR, FaccT...
  - USER MODELING AND USER-ADAPTED INTERACTION (UMUAI journal)
  - ACM Transactions on Recommender Systems (ToRS, introduced in 2021)

NSWI166 – Related Courses

- **NDBI021: User preferences and advanced recommending methods**

  Follow-up course in the winter semmester

  – More insights on what the user preferences are & how to learn them

  – Advanced concepts in Recommender Systems (Biases, multi-objective/multi-stakeholders, Fairness, Domain specific challenges)

Other Courses on Information Retrieval:

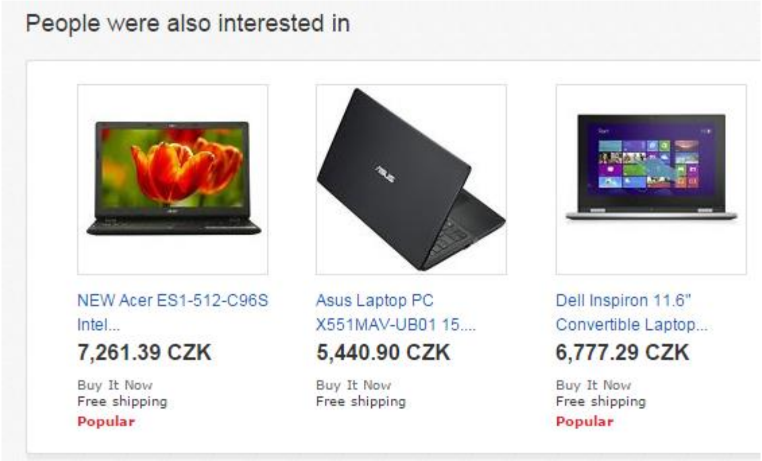
NDBI038: Searching the Web, NDBI045: Video Retrieval, NDBI034: Multimedia Retrieval (winter)

Questions?

RecSys „Hello worlds“ (if you leave now, please remember...)

Non-personalized

- no explicit notion of the current user (i.e., no UID)
- recommending based on aggregated data
    - and (optionally), current page context (item-based RS)



„Most (recently) popular items (in this category)“

„People who bought this also bought that“

„Similar objects to this one“

…

```
Select count(*) as num, item_i
      from coocurred_items
      where item_j = [current_item]
      group by item_i
      order by num
      limit 0, [num_of_displayed]
```

Caching / periodic evaluation for the sake of speed

RecSys „Hello worlds“ (if you leave now, please remember...)

- **The simplest personalized recommending algorithm**
- User-based nearest-neighbor collaborative filtering (UB-kNN)
  - Get most similar users to the current one.
    - Cosine similarity / Correlation of known ratings per user
  - Aggregate their feedback on not-yet-visited items.
    - Average / similarity-weighted average…
  - Recommend items with best aggregated feedack.
    - Top-k items with highest rating

RecSys „Hello worlds“ (if you leave now, please remember...)

- **Next lecture: 27.2.**
  - Why not to use User-based KNN
    - (no 90s nostalgia tolerated here)
  - What are better alternatives?
    - Why also do not use these?
    - What should I use then?
    - ???Why is everything so complicated and no answer is final???

Guess the title... ☺

?