

# NDBI021, Lecture 1

User preferences, 2/1 ZK+Z,

Wed 12:20 - 13:50 S8

Wed 14:00 - 15:30 SW2 (odd weeks)

<https://www.ksi.mff.cuni.cz/~peska/vyuka/ndbi021/2022/>



<https://ksi.mff.cuni.cz>

# NDBI021: User preferences

## Lectures

- ▶ User preferences (cca 8 lectures: 4x LP + 4x PV)
- ▶ Advanced topics from recommender systems (cca 5 lectures: LP)

## Labs

- ▶ TBA
  - ▶ Paper reports
  - ▶ Coding & analysis & results presentation

# NDBI021: User preferences

## Syllabus

### ► **User preferences**

- Introduction, motivation, challenges and use-cases of user preferences
- Modelling / expressing user preferences, types of user feedback, LMPM
- Learning user preferences, feedback interpretation, aggregating preferences & fuzzy logic
- Applications for user preferences, recommender systems, personalized search, challenge-response framework

### ► **Advanced topics from recommender systems**

- Fairness and proportionality in recommender systems
- Multicriterial optimization & evaluation in recommender systems
- Dynamic recommender systems: multi-armed bandits & reinforcement learning
- Unbiased evaluation, inverse propensity, feedback loops problem
- Deep learning for heterogeneous recommender systems

# NDBI021: User preferences

## Disclaimer

**Not about sandbox models**



~~We used exactly the same settings as those other guys, but significantly improved p@10 by 0.05%~~

# User preference

- ▶ **What is User Preference?**
- ▶ [IGI-global dictionary] A soft requirement provided by **users**, in addition to a query, to reflect a **wish**. It influences the querying process by causing some results to be favored.
  - ▶ Why/where to utilize user preferences (UP)?
  - ▶ How to define/express UP? How to model UP?
  - ▶ How to learn UP / interpret user's feedback?
  - ▶ What to do with UP?

# Related concepts

- ▶ (Implicit / explicit ) feedback
  - ▶ Typical means of getting user preferences
- ▶ Psychological profiling ([https://www.researchgate.net/publication/220737365\\_Towards\\_User\\_Psychological\\_Profile](https://www.researchgate.net/publication/220737365_Towards_User_Psychological_Profile) )
  - ▶ E.g. Sufficer vs. Maximizer (<https://www.wsj.com/articles/how-you-make-decisions-says-a-lot-about-how-happy-you-are-1412614997> )
  - ▶ Related area of user profiling
- ▶ Recommender systems
  - ▶ Typical means of utilizing user preferences
- ▶ Preference Elicitation
  - ▶ How to get (explicit) preferences (mainly) from new users

# Additional resources

▶ UMAP conference

- ▶ <https://www.um.org/umap2022/>

▶ UMUAI journal

- ▶ <http://www.umuai.org/>

▶ RecSys, SIGIR, ECIR, IUI, CHI...

- ▶ <https://sigir.org/sigir2022/>, <https://recsys.acm.org/>

## UMAP tracks 2022

- [Personalized Recommender Systems](#)
- [Adaptive Hypermedia, Semantic, and Social Web](#)
- [Intelligent User Interfaces](#)
- [Technology-Enhanced Adaptive Learning](#)
- [Fairness, Transparency, Accountability, and Privacy](#)
- [Personalization for Persuasive and Behavior Change Systems](#)
- [Virtual Assistants and Personalized Human-robot Interaction](#)
- [Research Methods and Reproducibility](#)

# Why / where to utilize user preferences?

## Why:

- ▶ **Information Overload** ([https://en.wikipedia.org/wiki/Information\\_overload](https://en.wikipedia.org/wiki/Information_overload))
  - ▶ Difficulty to effectively decide when one has too much information (options to choose from)
  - ▶ Filter available information according to user's preferences
- ▶ Because it may improve user's experience => satisfaction, loyalty, consumption...
  - ▶ => improve target metrics of the content provider
  - ▶ *Knowing your users allows you to serve them better => monetize them better*





# Why/where to utilize user preferences?

## Where & what to do:

- ▶ Limit ourselves to information retrieval (IR) domain
  - ▶ Within IR, it may be relevant throughout various domains
  - ▶ Recommender systems,
  - ▶ GUI personalizations (<https://evergreen.team/articles/ui-and-ux-personalization.html>)
  - ▶ Personalized search ([https://en.wikipedia.org/wiki/Personalized\\_search](https://en.wikipedia.org/wiki/Personalized_search))
- ...
- ▶ Know what you want to achieve (task is important)

Recent Automatic Recommendations

1-34 of 202 ( next ) titles | covers | shelf

Feb 2 Feb 2 Feb 2 Feb 2 Feb 2

IAIN M. BANKS THE HYDROGEN SONATA TERRY PRATCHETT JASPER FFORDE I WAS THERE NIGHT CIRCUS JOHN SCALZI

jaguar search

Conventional

jaguar search

Personalize

The image illustrates two search results for 'jaguar'. The 'Conventional' search shows a grid of images including a tree, a car, a red car, a blue car, a car wheel, and a leopard. The 'Personalize' search shows a grid of images including a tree, a leopard, a jaguar, a tiger, and a lion. A blue box labeled 'Personalize' has an arrow pointing to the personalized search results.

# How to define user preference model?

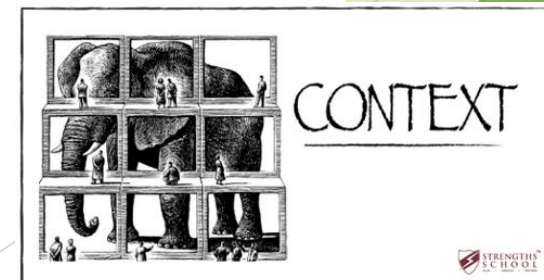
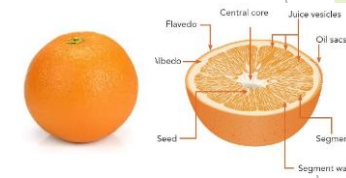
## Highly domain / task specific:

### ▶ What is the task?

- ▶ What is/are the entity/ies on which preferences can be expressed?
  - ▶ Objects? Their features? Groups of items?... Page layout?
- ▶ Who gives the preference (who is the user)?
  - ▶ An individual? A group of people? A cookies stored in a browser?
- ▶ What are expected connections among these entities (hypotheses)?
  - ▶ Some features are more important than others (presumably)
  - ▶ Preference of an object corresponds to an aggregation of its feature's preferences)
- ▶ Does user preference depend on some context/conditions?
  - ▶ I only watch romantic comedies with my wife...

### ▶ How to express preferences?

- ▶ Implicitly? Explicitly? A combination of both? Sth. else?



# What tools does the user have (to express preferences)?

- ▶ „Expressing by doing“ (implicit feedback)
- ▶ Rating/(dis)approving (explicit feedback)
- ▶ Filtering, searching
- ▶ Explicit comparison (A is better than B)
- ▶ Writing a review
  
- ▶ Did I forgot anything?

# How to collect user preferences?

- ▶ **Rating, filtering, comparison, reviews... via designated GUI**
  - ▶ How to store e.g. searching / filtering may be a bit tricky...
- ▶ **Implicit feedback**
  - ▶ **Server-side (limited expressibility)**
  - ▶ **Client-side (triggered JS events)**
  - ▶ **Beyond (eye tracking, other biometrics)**
    - ▶ Limited applicability (lab studies)
    - ▶ Can provide leads on interpretation of the previous two
- ▶ **Questionnaires, role playing**
  - ▶ Lab studies only (in most cases)
  - ▶ Can provide leads on interpretation of other collection methods

# What to do with user preferences?

(Again) task specific

- ▶ **Apply it in recommender systems**
  - ▶ Most prominent use-case
  - ▶ Recommend objects to interact with
  - ▶ Subcase: construct streams of objects (music, learning,...)
- ▶ **Modified search engine**
  - ▶ Results re-ranking
- ▶ **Modified page composition**
  - ▶ Prepend most interesting categories
  - ▶ Re-order / hide / show some widgets based on utilization,...

# How to model user preferences

What can be expressed & how to express it

# How to model user preferences?

- ▶ What is/are the entity/ies on which preferences can be expressed?
- ▶ Who gives the preference (who is the user)?
- ▶ What are expected connections among these entities (hypotheses)?
- ▶ Does user preference depend on some context/conditions?
- ▶ In what way do we allow users to express their preferences?

# How to model UP [some use-cases]

## Simple movies recommendation:

- ▶ **Task: help to discover what to watch tonight**
- ▶ **How to use UP: Collaborative recommendation of movies**
  - ▶ E.g. Select users similar to the current one, recommend what other users liked but the current one does not know.

## (Food) Recipes recommendation:

- ▶ **Task: help to decide what to cook**
  - ▶ **How to use UP: personalized searching**
    - ▶ Searching for recipes matching desired ingredients + users preferences
  - ▶ **How to use UP: front-page recommendation**
    - ▶ E.g. recommend recipes from user's favorite ingredients, similar (not too similar) to the ones liked in the past



# How to model UP [some use-cases]

## Group music recommendation:

- ▶ **Task: create a background music playlist for an evening with friends**
  - ▶ On top of Youtube / Spotify or similar service
- ▶ **How to utilize it: fairness-aware playlist construction**
  - ▶ (Endless) sequence of tracks, representing interests of all group members

# Points to consider in modeling phase

- ▶ **What entities are relevant?**
  - ▶ Are they sufficiently independent? Can we get enough data about them?
  - ▶ How much added value they can give us? Would it upset our users if they are not allowed to prefer them?
- ▶ **What (minimal) granularity to use?**
  - ▶ Primary target: cooking recipes
    - ▶ Recipes are composed from multiple ingredients (and other attributes)
      - ▶ Ingredients themselves has some attributes (e.g. quantity, type of preparation, quality, category)
        - ▶ **0.5kg** of **ground free-ranged chicken breasts** -> chicken breasts are a type of meat / chicken meat / light meat...
  - ▶ *Would the user like the recipe more/less if it would be meat-free?*
  - ▶ *Is it possible that he/she (dis)like chicken breasts?*
  - ▶ *Can you say that the recipe contains too much meat?*
  - ▶ *Is it even possible to rate the recipe based on ingredients?*
- ▶ **What would the user be willing to do/say/rate/...?**
  - ▶ Users tend not to rate things much
  - ▶ But they are sometimes willing to confirm/reject your beliefs
  - ▶ They do browsing, searching, buying etc.

# What entities are relevant?

## ▶ What entities are relevant?

- ▶ Are they sufficiently independent?
  - ▶ Is the user's preference towards some recipe a mere aggregation of his/her preference towards its ingredients?
  - ▶ Do you like the song just because it was written by... John Lennon?
- ▶ How much added value they can give us?
  - ▶ Is it usual that people like Offsprings, but dislike their Americana album?

## ▶ Who should decide?

- ▶ Domain expert
- ▶ Supported by data analysis
  - ▶ If you have some preliminary data:
    - ... and/or continuous testing
    - ... and/or dedicated lab study
  - ▶ Look for both common cases and outliers. If every person is an outlier in some scenario & you only solve common cases... Not good 😊



# What minimal granularity to chose?

► This is the feedback we have



## Pálivá dušená mrkev

Zeleninová příloha, ale klidně i hlavní chod pro vegetariány.

[ČÍST VÍCE](#)

60min | Přílohy | česká kuchyně | vařilo 5 lidí

★ 4.5  
16 hodnocení



## Klasické kuře na paprice

Velmi jednoduchý recept na kuře na paprice.

[ČÍST VÍCE](#)

75min | Omáčky | česká kuchyně | vařilo 170 lidí

★ 4.5  
1050 hodnocení



## Vepřové nudličky Gyros

Vepřové nudličky při smažení provoní vaši kuchyň řeckým kořením.

[ČÍST VÍCE](#)

30min | Hlavní chody | řecká kuchyně | vařilo 102 lidí

★ 4.6  
509 hodnocení



► Would he/she like this one?



## Hovězí maso na mrkvi

Hlavní chod z hovězího masa. Recept zvládne i kuchař začátečník.

[ČÍST VÍCE](#)

90min | Hlavní chody | česká kuchyně | vařilo 17 lidí

★ 4.6  
102 hodnocení

# What minimal granularity to chose?

► This is the feedback we have



## Pálivá dušená mrkev

Zeleninová příloha, ale klidně i hlavní chod pro vegetariány.

[ČÍST VÍCE](#) 60min | Přílohy | česká kuchyně | vařilo 5 lidí

★ 4.5  
16 hodnocení



## Klasické kuře na paprice

Velmi jednoduchý recept na kuře na paprice.

[ČÍST VÍCE](#) 75min | Omáčky | česká kuchyně | vařilo 170 lidí

★ 4.5  
1050 hodnocení



## Vepřové nudličky Gyros

Vepřové nudličky při smažení provoní vaši kuchyň řeckým kořením.

[ČÍST VÍCE](#) 30min | Hlavní chody | řecká kuchyně | vařilo 102 lidí

★ 4.6  
509 hodnocení



► Would he/she like this one?



## Hovězí maso na mrkvi

Hlavní chod z hovězího masa. Recept zvládne i kuchař začátečník.

[ČÍST VÍCE](#) 90min | Hlavní chody | česká kuchyně | vařilo 17 lidí

★ 4.6  
102 hodnocení

# What minimal granularity to chose?

► This is the feedback we have



## Pálivá dušená mrkev

Zeleninová příloha, ale klidně i hlavní chod pro vegetariány.

ČÍST VÍCE

60min | Přílohy | česká kuchyně | vařilo 5 lidí

★ 4.5  
16 hodnocení



## Klasické kuře na paprice

Velmi jednoduchý recept na kuře na paprice.

ČÍST VÍCE

75min | Omáčky | česká kuchyně | vařilo 170 lidí

★ 4.5  
1050 hodnocení



## Vepřové nudličky Gyros

Vepřové nudličky při smažení provoní vaši kuchyň řeckým kořením.

ČÍST VÍCE

30min | Hlavní chody | řecká kuchyně | vařilo 102 lidí

★ 4.6  
509 hodnocení



► Would he/she like this one?



## Hovězí maso na mrkvi

ČÍST VÍCE

90min | Hlavní chody | česká kuchyně | vařilo 17 lidí

★ 4.6  
102 hodnocení

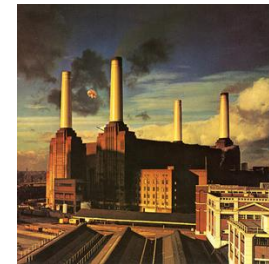
**Probably not, he/she seems to be a vegetarian**

What minimal granularity to chose?

**So, the finer granularity the better?**

# What minimal granularity to chose?

- ▶ Why do you like Lord of the Rings movie?
- ▶ What makes panna cotta so delicious?
- ▶ What is so good on [your favorite artists] songs?





# What minimal granularity to chose?

- ▶ Why do you like Lord of the Rings movie?
- ▶ What makes Panna cotta so delicious?
- ▶ What is so good on [your favorite artists] songs?
  
- ▶ *Is the description discriminative?*
  - ▶ *Can it be generalized to all other objects sharing the same feature?*
  - ▶ *Is there something missing in the picture?*
- ▶ *If you mentioned multiple features, what is their interplay?*
  - ▶ *And-like / or-like / additive / amplifying...*

# What minimal granularity to chose?

- ▶ Why do you like Lord of the Rings movie?
- ▶ What makes Panna cotta so delicious?
- ▶ What is so good on [your favorite artists] songs?
  
- ▶ *Do we have data for your choices?*
  - ▶ *How much meaningful are our attributes?*
  - ▶ *How much variety do we have?*
  - ▶ *(Hard to define) how much information is missing?*

# What minimal granularity to chose?

- ▶ Why do you like Lord of the Rings movie?
- ▶ What makes Panna cotta so delicious?
- ▶ What is so good on [your favorite artists] songs?
  
- ▶ *Also, are this the actual reasons, or just you justifying your emotions?*
  - ▶ <https://start.askwonder.com/insights/want-add-previous-request-made-few-years-ago-want-studies-support-idea-decisions-c22vygpap>
  - ▶ [https://en.wikipedia.org/wiki/Motivated\\_reasoning](https://en.wikipedia.org/wiki/Motivated_reasoning)
  - ▶ Ninety percent of human decisions are made based on emotions. Humans use logic to justify their actions to themselves and others
  - ▶ "Common emotional decisions may use some logic, but the main driving force is emotion, which either overrides logic or uses a pseudo-logic to support emotional choices (this is extremely common). Another common use of emotion in the decision is to start with logic and then use emotion in the final choice."

# What minimal granularity to chose?

To sum-up

- ▶ It may be difficult for humans to express their preference on sub-object level
  - ▶ Some hints are possible, but the information is rarely complete
- ▶ Our data does not have to correspond to the user's vocabulary
  - ▶ Some features are missing
  - ▶ Some user's preferences are hard to be transferred to attributes  
(The movie is too generic)
- ▶ Quite often, human's decisions are not driven logically, but emotionally
  - ▶ Any logical clues they give us does not have to be applicable for the next case

# What minimal granularity to chose?

One possible solution:

- ▶ Collect feedback on object-level, infer sub-object preferences
  - ▶ Latent or explicit
  - ▶ Occasionally ask, whether your explicit preferences are correct
    - ▶ E.g. Use latent model, but „guard“ it via regularization to the explicit one. Iteratively update if you receive confirmation on your explicit model. Or mixed latent + explicit model.  
<https://www.frontiersin.org/articles/10.3389/fdata.2021.778417/full>  
<https://dl.acm.org/doi/10.1145/2043932.2043979>
  - ▶ Confirmation is often easier than definition + can be done in a less intrusive way
    - ▶ Based on your recipe ratings, we think you are a vegetarian. Is that true?  
[Yes | NO | No, but I do not eat meat much | Irrelevant] What will happen >>  
  
If you click on Yes, we will promote more vegetarian meals in your future searches. You can always turn this feature off in your user profile.
  - ▶ User controll in RS ([https://web-ainf.aau.at/pub/jannach/files/BOOK\\_CHAPTER\\_PERSONALIZED\\_HCI\\_2019.pdf](https://web-ainf.aau.at/pub/jannach/files/BOOK_CHAPTER_PERSONALIZED_HCI_2019.pdf))
- ▶ Requires rather lot of per-user data to learn this reliably

# What minimal granularity to chose?

Be careful while generalizing from inferred sub-object preferences

- ▶ Do you like action movies?
- ▶ Do you like food with salt?
- ▶ Do you like movies with David Aston?
  - ▶ Yes, but not all of them
  - ▶ Yes, but not because of this
- ▶ ...who is David Aston?

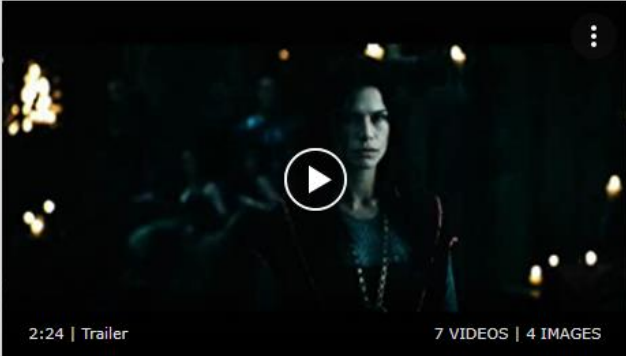

# What minimal granularity to chose?

Be careful while generalizing from inferred sub-object preferences

- ▶ Do you like action movies?
- ▶ Do you like food with salt?
- ▶ Do you like movies with David Aston?
  - ▶ Yes, but not all of them
  - ▶ Yes, but not because of this
- ▶ ...who is David Aston?

David Aston (1) SEE RANK


Actor



2:24 | Trailer 7 VIDEOS | 4 IMAGES



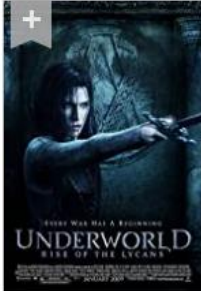

David Aston is an actor, known for *Matrix* (1999), *Pán prstenů: Návrat krále* (2003) and *Underworld 3: Vzpoura Lycanů* (2009). [See full bio »](#)

### Photos



[4 photos](#) | [7 videos](#) »

### Known For

 <p>Matrix Rhineheart (1999)</p>	 <p>Pán prstenů: Návrat krále Gondorian Soldier 3 (2003)</p>	 <p>Underworld 3: Vzpoura L... Coloman (2009)</p>	 <p>Seekers Chris / Chris Youngman ... (1986)</p>
---	---	--	--

# What minimal granularity to chose?

Be careful while generalizing from inferred sub-object preferences

- ▶ Yes, but not all of them
  - ▶ Should be fine (other features to clarify), but we need to ask questions correctly
- ▶ Yes, but not because of this
  - ▶ Loss of trust („*the system does not understand what is important*“)
  - ▶ Some model of global or per-object feature importance
    - ▶ Only confirm the inferred preference on the important ones
    - ▶ Do not base your decision on the less relevant much



What would the user be willing to do?

**How does the industry feel about that?**


# What would the user be willing to do?

#EnnioMorricone #EnnioMorriconeMusic #SpaghettiWesternMusic

Ennio Morricone - Sergio Leone Greatest Western Music of All Time (Remastered HQ Audio)


16 864 622 zhlédnutí • 25. 4. 2018 • Maestro Ennio Morricone and his timeless m Zobrazit více

👍 148 tis.    🗨️ Nelíbí se    ➦ Sdílet    ⋮ Uložit    ...

 **Ennio Morricone** 🎵  
479 tis. odběratelů

**ODEBÍRAT**

**Komentáře**  
6,8 tis.

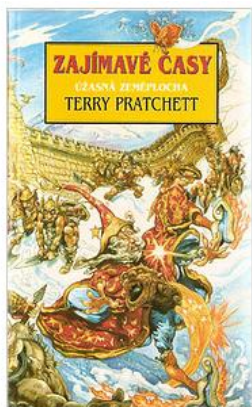
 In 100 years Ennio Morricone will be named together with Mozart, Bach, Puccini, Verdi and other great composers of the past. I'm sure!

6 849 komentářů

☰ ŘADIT PODLE



Přidejte veřejný komentář...



Member [Ipeska](#)

Review Not reviewed (edit)

Collections [Your library](#)

Rating ★★★★☆

Tags fantasy (edit)

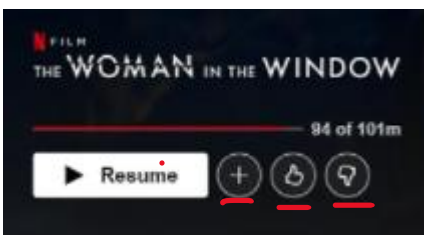
## Terminátor 2: Den zúčtování

Original title: Terminator 2: Judgment Day  
1991 · U · 2h 17m

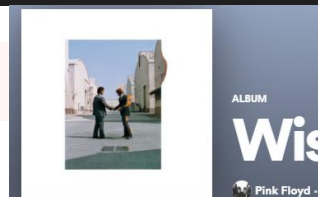
IMDb RATING  
★ 8.5/10  
1.1M

YOUR RATING  
★ Rate

POPULARITY  
📈 354 - 19



❤️ Líbí se 215    💬 Diskuze 169

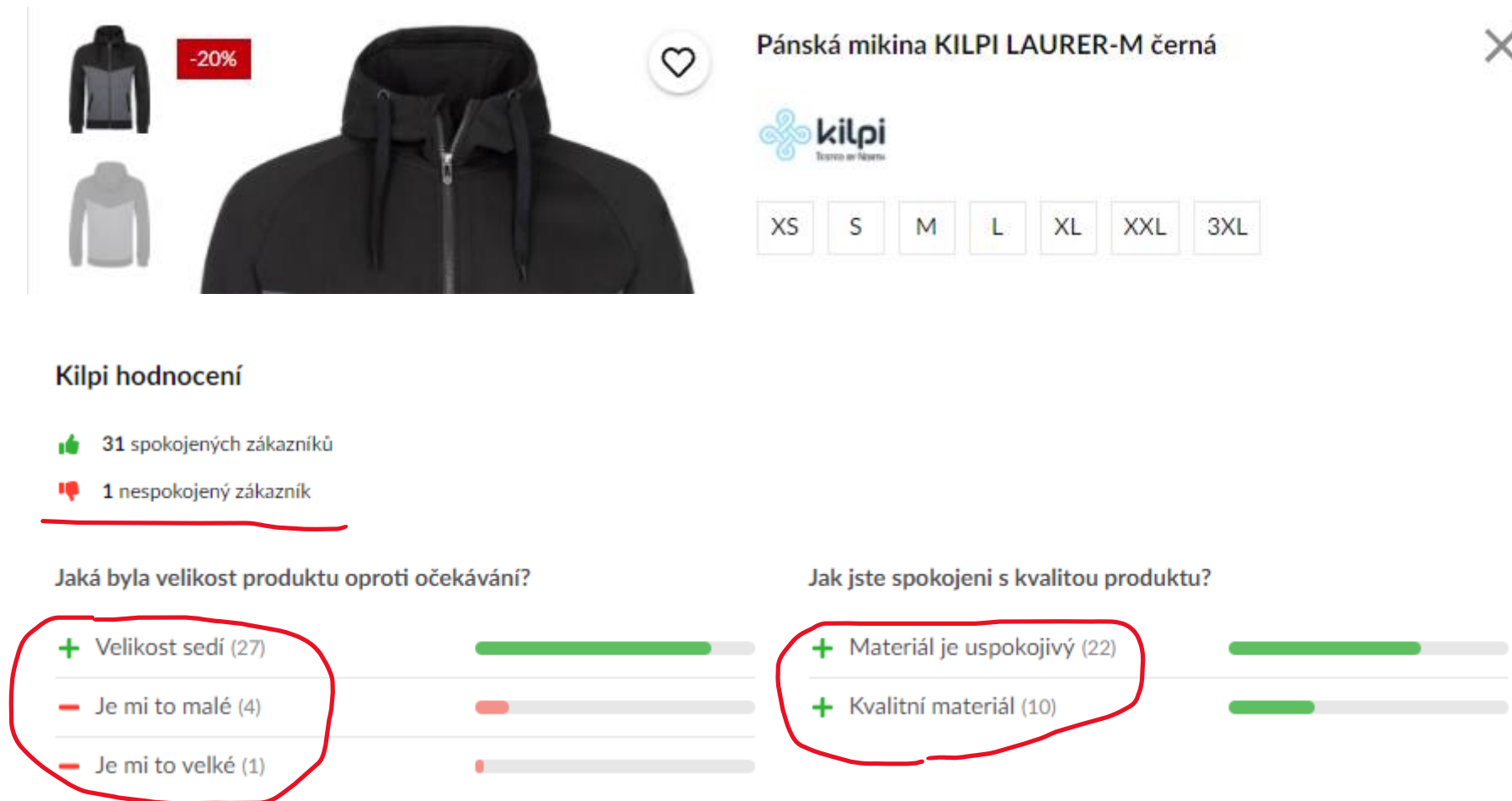


SDÍLEJTE ČLÁNEK    📧    🐦



Artist, album & track

# What would the user be willing to do?



The screenshot shows a product page for a black hoodie. The product name is "Pánská mikina KILPI LAURER-M černá". The brand logo "kilpi" is visible. Size options are XS, S, M, L, XL, XXL, and 3XL. The product has a -20% discount. Customer feedback is shown below the product image.

**Kilpi hodnocení**

👍 31 spokojených zákazníků  
👎 1 nespokojený zákazník

**Jaká byla velikost produktu oproti očekávání?**

- + Velikost sedí (27)
- Je mi to malé (4)
- Je mi to velké (1)

**Jak jste spokojeni s kvalitou produktu?**

- + Materiál je uspokojivý (22)
- + Kvalitní materiál (10)

# What would the user be willing to do?

Apartmány

## Ubytování u Kubů



Rezervujte si svůj pobyt v apartmánu

K dispozici je online check-in

Janáčkova 134/5, Jablonec nad Nisou, 466 06, Česká republika –

Rozdíl v ceně vyrovnáme

Skvělá lokalita – [ukázat mapu](#)



### Kategorie:

Personál



Zařízení ↓



Cistota ↓



Pohodlí



Poměr ceny a kvality



Lokalita



WiFi zdarma ↑



# What would the user be willing to do?




Miroslav, Praha 8 

Hodnoceno 22.10.2020, varianta 32" ViewSonic VX3276-2K-MHD

★★★★★  Ověřený nákup

- + velikost
- + obraz
- + vzhled
- + cena

- zvuk z reproduktoru

dejav 

Hodnoceno 25.09.2019, varianta 32" ViewSonic VX3276-2K-MHD

★★★★★

Dlouho jsem poohlížel po levném a velkém monitoru. Tento monitor je v nízké cenové relaci a jak se říká za málo peněz hodně muziky

- + Cena
- + Velikost
- + Kvalita
- + IPS

- zatím žádné

# What would the user be willing to do?

Some users are sometimes willing to

- ▶ **Provide ratings**
  - ▶ Sometimes aspect-based ratings (mostly pre-defined, widely recognized categories)
    - ▶ Does not have to correspond to object's attributes directly
  - ▶ Write review
- ▶ Add object to some list / organize favorite objects / provide tags for them
- ▶ Share items
  
- ▶ Is this frequent enough so we can infer preferences of individual users?

# What would the user be willing to do?

Cestuji pracovně

**Hledat podle:**  
**Váš rozpočet (na noc)**  
 Nastavit vlastní rozpočet  
   
**Populární filtry**  
 Snídaně v ceně 401  
 Lázně a wellness 86  
 Krytý bazén 27  
 Vlastní koupelna 868  
 Parkoviště 727  
 Méně než 1 km 272  
Vzdálenost od centra destinace Praha  
 Fantastické: 9 a více 279  
Na základě hodnocení hostů  
 4 hvězdičky 376  
**Hledejte ubytování podle názvu**  
  
**Zdraví a bezpečnost**  
 Ubytování, která zavedla zdravotní a bezpečnostní

**Cena**  
 -   
  
**Stav zboží**  
 Vše  
 Jen nové  
 Jen rozbalené, zánovní, použité  
**Značka**  
 Ratikon (7)  
 Leader Fox (4)  
 Sava (40)  
 Cycleman (4)  
**Typ kola**  
 Horské (43)  
 Trekingové (6)  
 Městské (8)  
 Elektrokolo (48)  
**Určeno pro**

**55 položek**  
       
**Úroveň hotelu**  
 ★  
 ★★  
 ★★★  
 ★★★★  
 ★★★★★  
**Cena**  
   
Minimální cena Maximální cena  
**Typ dovolené**  
**Vybavení hotelu**  
**Sport/zábava**  
**Vybavení pokoje**  
**Vzdálenost od aquaparku**  
**Doba transferu z letiště**

**OUTDOOR**  

- BOTY
- BATOHY
  - Pánské batohy
  - Dámské batohy
  - Dětské batohy
  - Cestovní tašky a duffle
  - Ledvinky, taštičky a peněženky
  - Dětská nosítka
  - Doplňky k batohům
- STANY
- SPACÁKY A KARIMATKY
- SVÍTILNY

# What would the user be willing to do?

## Most users do:

- ▶ **Filter content manually**
  - ▶ Browse categories
  - ▶ Apply facet search
    - ▶ Mostly direct mapping to object's attributes
  - ▶ Use fulltext search
  - ▶ Can be utilized in the construction of attribute-level preferences
    - ▶ Beware of long-term preferences vs. short-term goals

## All users do:

- ▶ **Evaluate & consume content:**
  - ▶ Browse items, open details, read content, play, purchase,...
  - ▶ Preferences based on implicit feedback



# How to model UP

## Simple movies recommendation:

- ▶ **Task: discover what to watch tonight**
- ▶ How to use UP: Collaborative recommendation of movies

## Basic model of UP:

- ▶ Preferences on movies (rating, watching)
  - ▶ If insufficient data: opening movie details, top search results

## Enhancements:

- ▶ Learned (confirmed) preferences towards genres (multiple confirmation, enough data)
- ▶ Learned (confirmed) preferences towards other named entities (actor, director)
- ▶ List-wise preferences (Y was selected from results of XYWZ)
  - ▶ Remember impressions, not just usage

# How to model UP

## (Food) Recipes recommendation:

- ▶ **Task: help to decide what to cook**
  - ▶ How to use UP: **personalized searching, front-page recommendation**

## Basic model of UP:

- ▶ Preferences on recipes (likes, add to list, reading sufficiently long)
- ▶ Preferences on ingredients (search count, contained in preferred recipes, confirmation?)
  - ▶ Ingredients granularity?

## Enhancements:

- ▶ Learned preferences towards tags & attributes
  - ▶ Verify on a well-known subset of users (RecSys OPS)
- ▶ Best out of similar choices
  - ▶ Which goulash does the user prefer? Would that say something more generic about him/her?
- ▶ Should we allow users to further refine recommendations?
  - ▶ Faceted recommendations  
([https://www.researchgate.net/publication/301321425\\_FeRoSA\\_A\\_Faceted\\_Recommendation\\_System\\_for\\_Scientific\\_Articles](https://www.researchgate.net/publication/301321425_FeRoSA_A_Faceted_Recommendation_System_for_Scientific_Articles))

# How to model UP

## Group music recommendation:

- ▶ **Task: create a background music playlist for an evening with friends**
- ▶ How to utilize it: fairness-aware playlist construction
  
- ▶ Individual preference
  - ▶ Track -> Album -> Artist (playcount, play from search, likes)
  - ▶ Maybe, preferred sequences (low-level audio analysis, but probably not for individual users)
  
- ▶ Group preferences
  - ▶ Playlist modifications

# Preference Elicitation

The background features abstract, overlapping geometric shapes in various shades of green, ranging from light lime to dark forest green. These shapes are primarily located on the right side of the slide, creating a modern, layered effect. The text 'Preference Elicitation' is positioned on the left side of the slide in a clean, sans-serif font.

# Preference Elicitation

- ▶ [WIKI] Preference elicitation refers to the problem of developing a [decision support system](#) capable of generating [recommendations](#) to a user, thus assisting in decision making. It is important for such a system to model user's preferences accurately, find hidden preferences and avoid redundancy.
- ▶ Not really a definition
- ▶ The process of collecting user preferences to support decision making systems
  - ▶ Often considered w.r.t. restricted meaning of initial preference elicitation
  - ▶ Usually restricted to explicit feedback

Traditional methods (2004):

<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.319.8057&rep=rep1&type=pdf>

# Preference Elicitation

- ▶ Additive independence of preferences
  - ▶ Preferences of items is a function of it's features preferences (wAVG)

**Mutual Preferential Independence**: The attributes  $X = \{x_1, \dots, x_n\}$  are mutually preferentially independent if every subset  $Y$  of  $X$  is preferentially independent of its complementary set.

**Theorem of Additive Value Function**: Given attributes  $X = \{x_1, \dots, x_n\}$ ,  $n \geq 3$ , an additive value function  $v(x_1, \dots, x_n) = \sum_{i=1}^n \lambda_i v_i(x_i)$  (where  $v$  and  $v_i$  are scaled from zero to one, and

$\sum_{i=1}^n \lambda_i = 1, \lambda_i > 0$ ) exists if and only if the attributes are mutually preferentially independence.

**Additive Independence**: If the value function can be wrote as additive model, namely the condition of mutually preferentially independence is met, the attributes are said to be additive independent.

# Preference Elicitation

## ► Additive independence of preferences

**Mutual Preferential Independence:** The attributes  $X = \{x_1, \dots, x_n\}$  are mutually preferentially independent if every subset  $Y$  of  $X$  is preferentially independent of its complementary set.

**Theorem of Additive Value Function:** Given attributes  $X = \{x_1, \dots, x_n\}$ ,  $n \geq 3$ , an additive value function  $v(x_1, \dots, x_n) = \sum_{i=1}^n \lambda_i v_i(x_i)$  (where  $v$  and  $v_i$  are scaled from zero to one, and

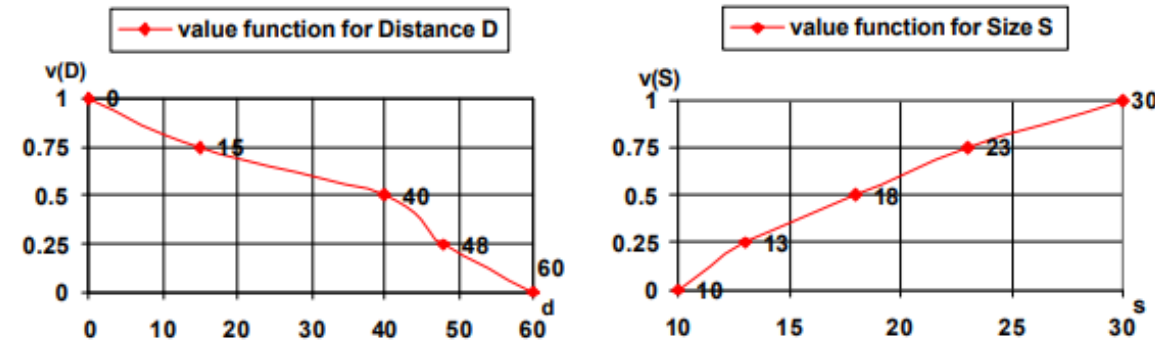
$\sum_{i=1}^n \lambda_i = 1, \lambda_i > 0$ ) exists if and only if the attributes are mutually preferentially independent.

**Additive Independence:** If the value function can be written as an additive model, namely the condition of mutual preferential independence is met, the attributes are said to be additive independent.

## ► Similar as LPM - only value functions do not have to be linear

Question	Hypothesized answer
1. Suppose you are at Size=20. Would you pay more of Size to change Distance from 60 to 30 or 30 to 0?	I would pay more to go from 60 to 30.
2. More to go from 60 to 50 or 50 to 0?	More to go from 50 to 0.
3. Give me a value, $d'$ , say, such that you would give up the same in Size to go from 60 to $d'$ as from $d'$ to 0.	About $x' = 40$
4. In our language, 40 is the midvalue point between 0 and 60. We label 40 by $d_{.5}$ . What is your midvalue point between 0 and 40?	Let's say 15, I'd pay the same to go from 40 to 15 as 15 to 0.
5. In that case $d_{.75} = 15$ . What is your midvalue point between 40 and 60?	Oh, about 48
6. This means that $d_{.25} = 48$ . Does 40 seem like a good midvalue between 15 and 48?	Sure
7. Now let's turn to the Size value. What is the midvalue point between 10 and 30?	Say, 18.
8. The midvalue between 18 and 30?	Say, 23.
9. The midvalue between 10 and 18?	13.

Then we can plot these few points and fit in the curves of  $v_D$  (distance) and  $v_S$  (size).



# Preference Elicitation

- ▶ Knowledge-based RS with preference elicitation
  - ▶ Start either with known example
  - ▶ Or initial search

The image shows a RentMe interface with a listing for a 2-bedroom apartment in Ukrainian Village. The listing includes a description, phone number, and price. Below the listing, there are two sections for preference elicitation. The first section, titled "This apartment is OK, but make it...", has four buttons: "bigger", "cheaper", "nicer", and "safer". The second section, titled "This neighborhood could be more...", has three buttons: "convenient", "conservative", and "dynamic". To the right, there are two more listings. The first listing is titled "These apartments have a cheaper rent." and has a "Yes, but..." button. The second listing is titled "These apartments are cheaper, but are in other neighborhoods." and has a "Yes, but..." button. Both listings also have an "Add to list" button.

Phone: 312-489-1554	2-bedrooms	\$600	60622 (West Town Backtown)
---------------------	------------	-------	----------------------------------

UKRAINIAN VILLAGE TWO bedroom w/hat garden apartment. Lt. Enclosed, hv/d, excellent security, forced air, lots of closets, laundry in building. Garage space included. Dogs OK. Available immediately. \$600/mo. 312-489-1554.

This apartment is OK, but make it...

bigger cheaper nicer safer

This neighborhood could be more...

convenient conservative dynamic

These apartments have a cheaper rent.

Phone: 378-6064	2-bedrooms	\$520	60622 (West Town Backtown)
-----------------	------------	-------	----------------------------------

UKRAINIAN VILLAGE SPECIAL 2 bedroom. Hard wood floors, pocket doors, tin ceiling, pantry. Storage and parking included. Very sunny. \$520. Available immediately. 378-6064.

Yes, but...

Add to list

These apartments are cheaper, but are in other neighborhoods.

Phone: 312-338-0199	2-bedrooms	\$510	60626
---------------------	------------	-------	-------

VERY COZY ROGERS Park two bedrooms (Aristo Dancer). Hardwood floors, armoire, completely renovated kitchen, large closets, updated bath, freshly painted, cable ready, small deck, 24 hour maintenance, laundry, storage. \$510 includes heat. Phone 312-338-0199 or JED 708-678-5512.

Yes, but...

Add to list

Figure 2 Tweaking an apartment in RentMe



# Preference Elicitation

- ▶ <https://www.frontiersin.org/articles/10.3389/frobt.2017.00071/full> (Constructive pref. Elicitation)
  - ▶ There exist many types of queries, like lotteries, pairwise or setwise rankings, improvements, which all share the goal of being easy to answer to and as informative as possible.
    - ▶ Choice set feedback
    - ▶ Coactive feedback (how to slightly improve a solution? - can be done from implicit feedback) <https://www.jair.org/index.php/jair/article/view/10939>
    - ▶ Example critiquing
  - ▶ Queries involving comparisons and rankings have come to be predominant in the literature with respect to quantitative evaluations.
  - ▶ Indeed, users are typically more confident in providing qualitative judgments like “I prefer configuration  $y$  over  $y'$ ” than in specifying how much they prefer  $y$  over  $y'$  ([Conitzer, 2009](#); [Carson and Louviere, 2011](#)).

# Preference Elicitation

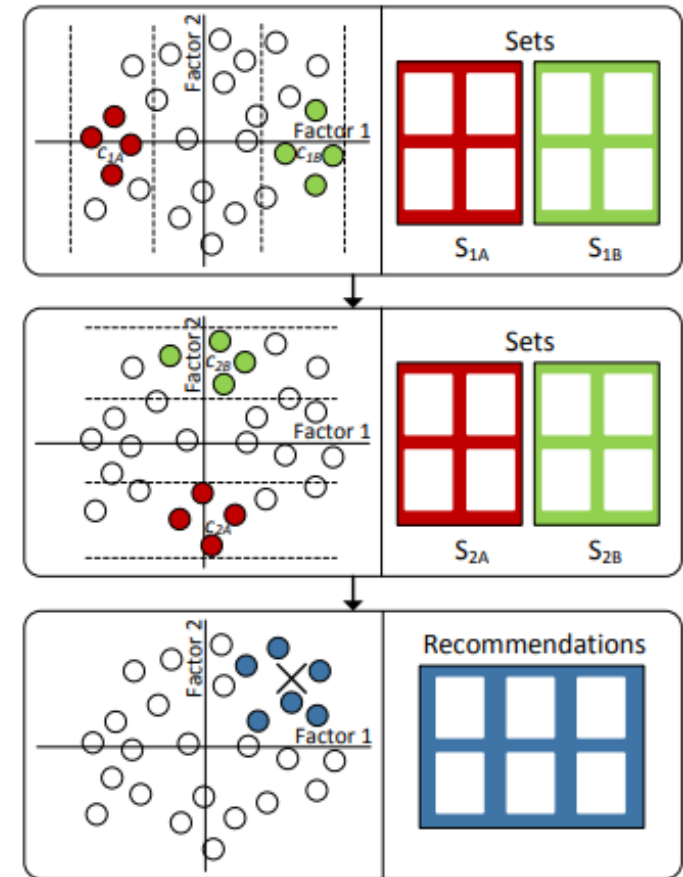
- ▶ <https://dl.acm.org/doi/pdf/10.1145/2792838.2796554> (Healthy recipes recommendation)
  - ▶ What was the main cause of your decision?
    - ▶ Video:  
<https://onedrive.live.com/?authkey=%21ALYePnW0fOCHOUQ&cid=60DC0855E37985A6&id=60DC0855E37985A6%2149418&parId=60DC0855E37985A6%2149101&o=OneUp>
    - ▶ Relatively simple tag-based approach

# Preference Elicitation

- ▶ Choice-based preference elicitation for collaborative filtering recommender systems

<https://dl.acm.org/doi/10.1145/2556288.2557069>

- ▶ Not based on meta-data, but latent factors
  - ▶ „The basic idea behind our approach is, thus, to use latent item features derived from the rating matrix and request preferences for sets of similar items instead of single items.“
  - ▶ „Since the number of interaction steps needed should be minimized, we developed a technique based on latent factors to achieve a maximum information gain with each choice.“



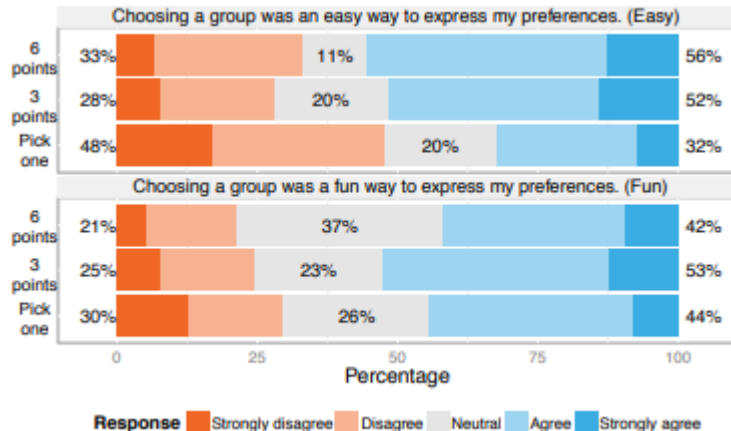
**Figure 2.** For each factor  $f$  taken into account, two sets of movies  $S_{fA}$  and  $S_{fB}$  are presented to the user. One set shows movies with low factor values, the other movies with high factor values. The user selects one of these sets (or indicates that he/she doesn't care). After a defined number of steps, a set of recommendations is computed.

# Preference Elicitation

## Using Groups of Items for Preference Elicitation in Recommender Systems

<https://dl.acm.org/doi/pdf/10.1145/2675133.2675210>

- ▶ New users can begin by expressing their preferences for groups of items
  - ▶ Utilize clustering to generate groups
  - ▶ Based only on movie ratings
  - ▶ For each cluster: select tags, then select best matching movies
  - ▶ Get avg. ratings of users with similar cluster prefs.



movielens

What kind of movie fan are you? Distribute 6 points among the groups of movies below to represent your preferences. MovieLens will then recommend movies personalized to your selection.

Next  Remaining points:

**Representative tags**: courage, earnest, touching

**Representative movies**: Braveheart, Apollo 13, Million Dollar Baby

**Points assigned to cluster**: 2

**Representative tags**: dark humor, enigmatic, masterpiece,

**Representative movies**: Fargo, The Godfather, A Clockwork Orange

**Representative tags**: based on a comic, dark hero, superhero

**Representative movies**: Batman Begins, Iron Man, The Avengers

**Representative tags**: computer game, explosions, sci-fi

**Representative movies**: MIB, Men in Black, The Matrix, I, Robot

For each movie group, we first pick the top-three tags that both *uniquely describe* and are *highly relevant* to the group. Therefore, we define the measure of tag uniqueness as Equation 1 and tag relevance as Equation 2. We pick the three tags with the highest multiplication of uniqueness and relevance. (Multiplication is used to handle different scales of the two metrics.)

$$unique(t, c) = \frac{rel(t, c)}{\sum_{c_i \in C} rel(t, c_i)} \quad (1)$$

$$relevance(t, c) = \frac{rel(t, c)}{\sum_{t_i \in T_c} rel(t_i, c)} \quad (2)$$

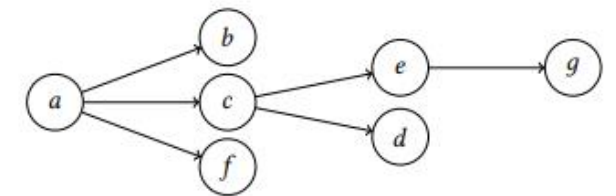
where  $t$  denotes one of the tags  $T_c$  that appears in cluster  $c$ , and  $C$  denotes all the clusters. Note that  $rel(t, c)$  is the aggregated relevance of tags  $t$  to all movies in cluster  $c$ . In our implementation, we use relevance between a tag and a movie generated from the Tag Genome [31],

# Preference Elicitation

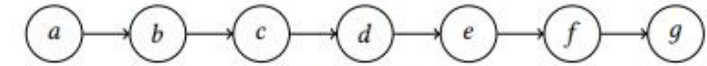
## Ordered Preference Elicitation Strategies for Supporting Multi-Objective Decision Making

<https://arxiv.org/pdf/1802.07606.pdf>

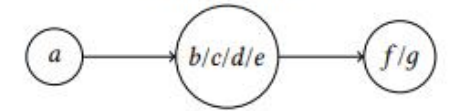
- ▶ Utilize full ranking of items
  - ▶ User starts with two items, then iteratively place one more item at each step
  - ▶ How to select what to ask?
    - ▶ Gaussian process (model mean and variance for each datapoint) (<https://ebonilla.github.io/gaussianprocesses/>, <https://github.com/chariff/GPro> )
    - ▶ Expected improvement acquisition function ([https://www.csd.uwo.ca/~dlizotte/publications/lizotte\\_phd\\_thesis.pdf](https://www.csd.uwo.ca/~dlizotte/publications/lizotte_phd_thesis.pdf))



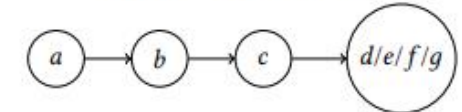
(a) Pairwise Comparisons



(b) Ranking



(c) Clustering (2 clusters)



(d) Top-Rank (top 3)

**Figure 2: Possible outcomes of different query types for items  $a-g$ , with utilities  $u(a) > \dots > u(g)$ . The arrows represent the preference information expressed by the user (preferred  $\rightarrow$  unfavoured). Different elicitation strategies lead to different orderings: full ranking returns a total ordering (b); the other query types typically lead to partial orderings.**

