RS

recommender systems

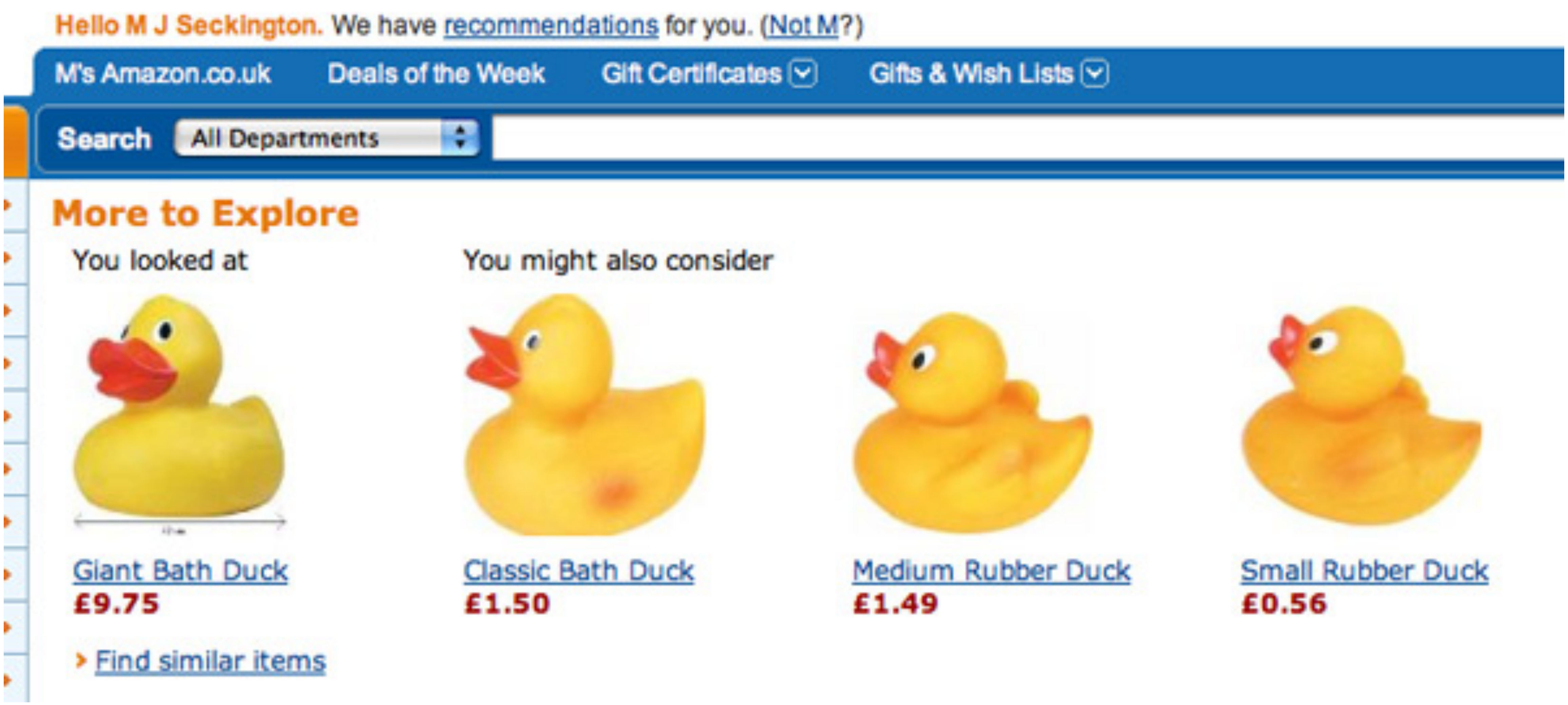
PG

28.9.2017

Source of background: www.wallpaperswide.com

datalys

What is recommender?



datalys

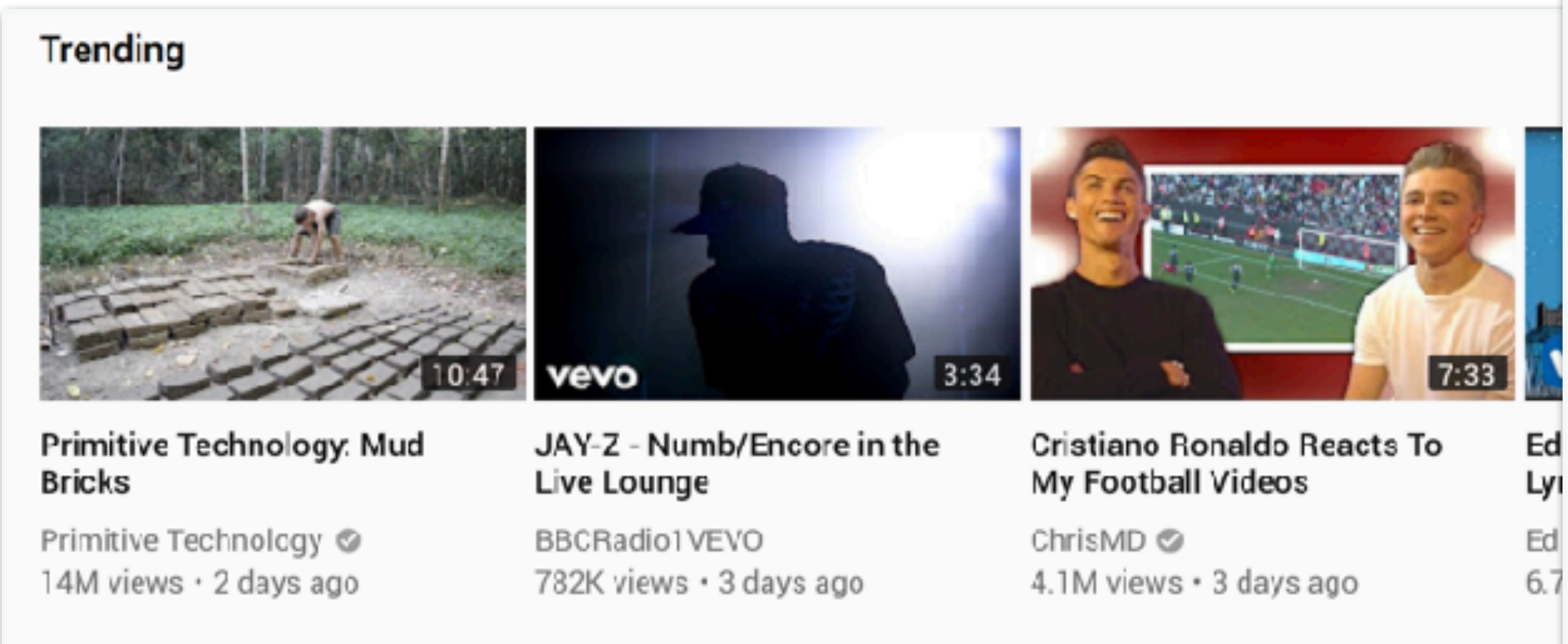
Why we need them?

Plenty of information available “out there”.

datalys

Popularity != ideal

Basic - recommend the most popular.



NAJČÍTANEJŠIE NA SME

4 hod 24 hod 7 dní SME+ EN

1. Nie sme ovce, Gašpar preč, znelo na pochode proti korupcii (minúta po minúte) VIDEO 21 676
2. Štvrtý titul? O rok naň Sagan môže zabudnúť 14 270
3. Sagan či Hantuchová sa budú plateniu daní na Slovensku vyhýbať ťažšie 11 927
4. Českého turistu našli v Tatrách mŕtveho. Fatálne podcenil výstroj a počasie FOTO 11 258
5. Pozrite si záver MS z Bergenu, ktorý nebol v televízii VIDEO 10 298
6. Päťdesiatnici a stres: Po päťdesiatke začne stres zabíjať 7 565
7. Ešte v marci robil v obleku kľuky. Železná vôľa urobila z Jana Třísku Američana VIDEO 7 408
8. Nový vynález môže výrazne zlepšiť opravu výtlkov VIDEO 7 280

datalys

1. Darebák jedna: Star Wars Story (2016)
2. Finding Dory (2016)
3. Captain America: Civil War (2016)
4. The Secret Life of Pets (2016)
5. The Jungle Book (2016)
6. Deadpool (2016)
7. Zootopia (2016)
8. Batman v Superman: Dawn of Justice (2016)
9. Jednotka samovrahov (2016)
10. Sing (2016)

IMDb - Top 10 US Box office 2016

Rank	Rating	Title
1.	98%	Moonlight (2016)
2.	98%	Zootopia (2016)
3.	94%	Arrival (2016)
4.	96%	Manchester by the Sea (2016)
5.	97%	Hell or High Water (2016)
6.	95%	The Jungle Book (2016)
7.	92%	La La Land (2016)
8.	94%	Finding Dory (2016)
9.	96%	Moana (2016)
10.	98%	Love & Friendship (2016)

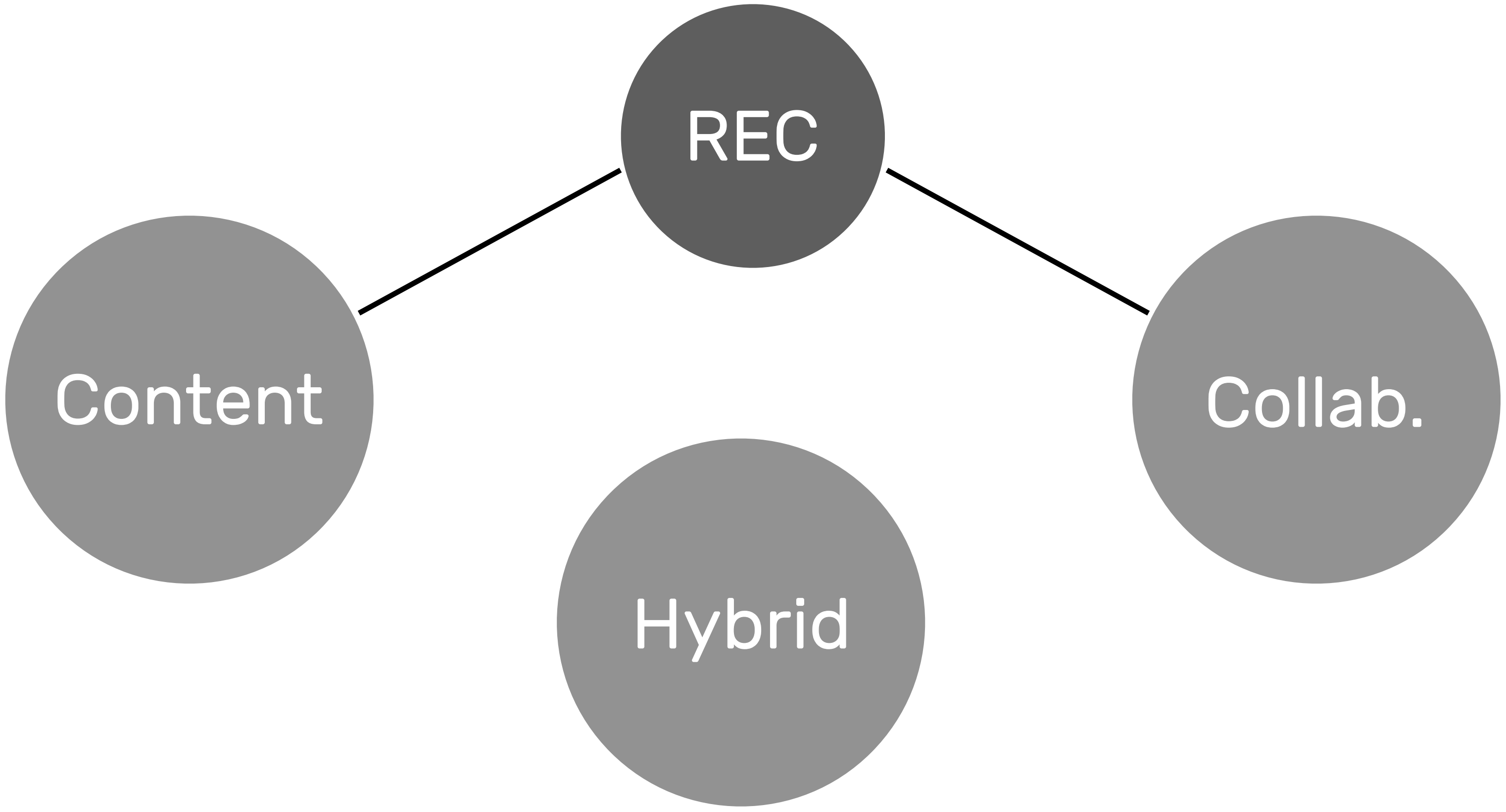
Rotten Tomatoes - Top 10 (rating)

Personalization

Tailor recommender to *user preferences*.

How?

Recommendation techniques

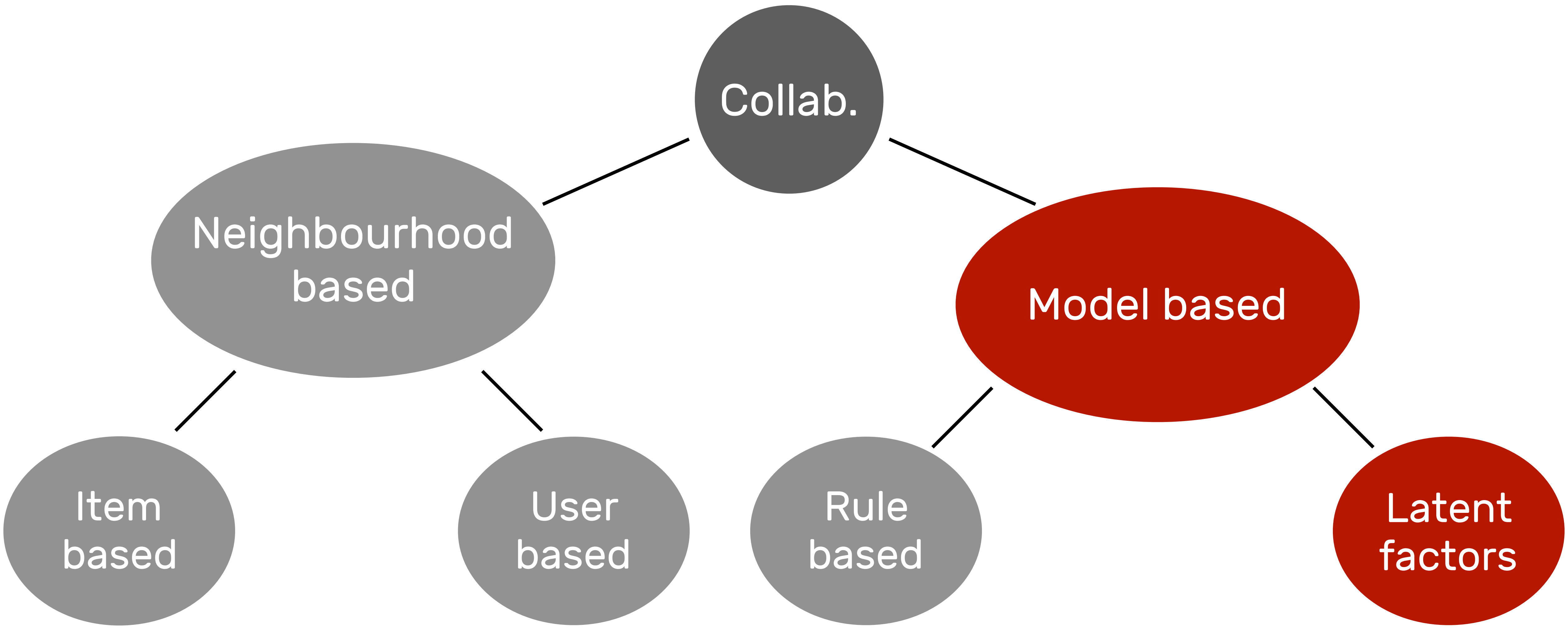


What about data?

Implicit feedback
clicks, purchases, plays

Explicit feedback
ratings, likes, reviews

Recommendation techniques

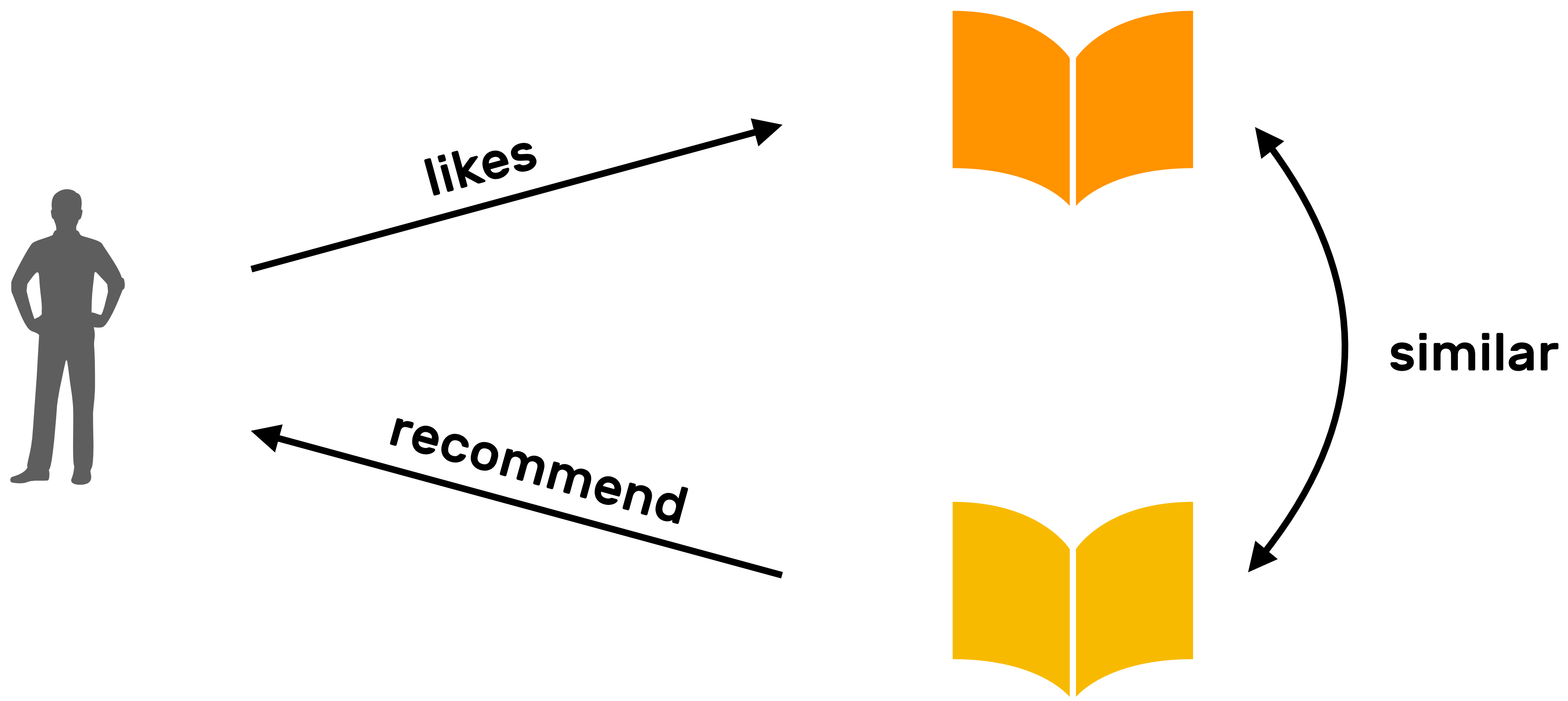


Matrix Factorization
Neural Networks

Content-based

Utilize (textual) characteristics of items.

News, movies, books, …



datalys

Collaborative filtering

Based on the items that users with the similar taste previously liked.

Two types:

- neighbourhood-based,
- model-based.

Problem: cold-start, similarity

Neighborhood-based

Focuses on the similarity between users or items.

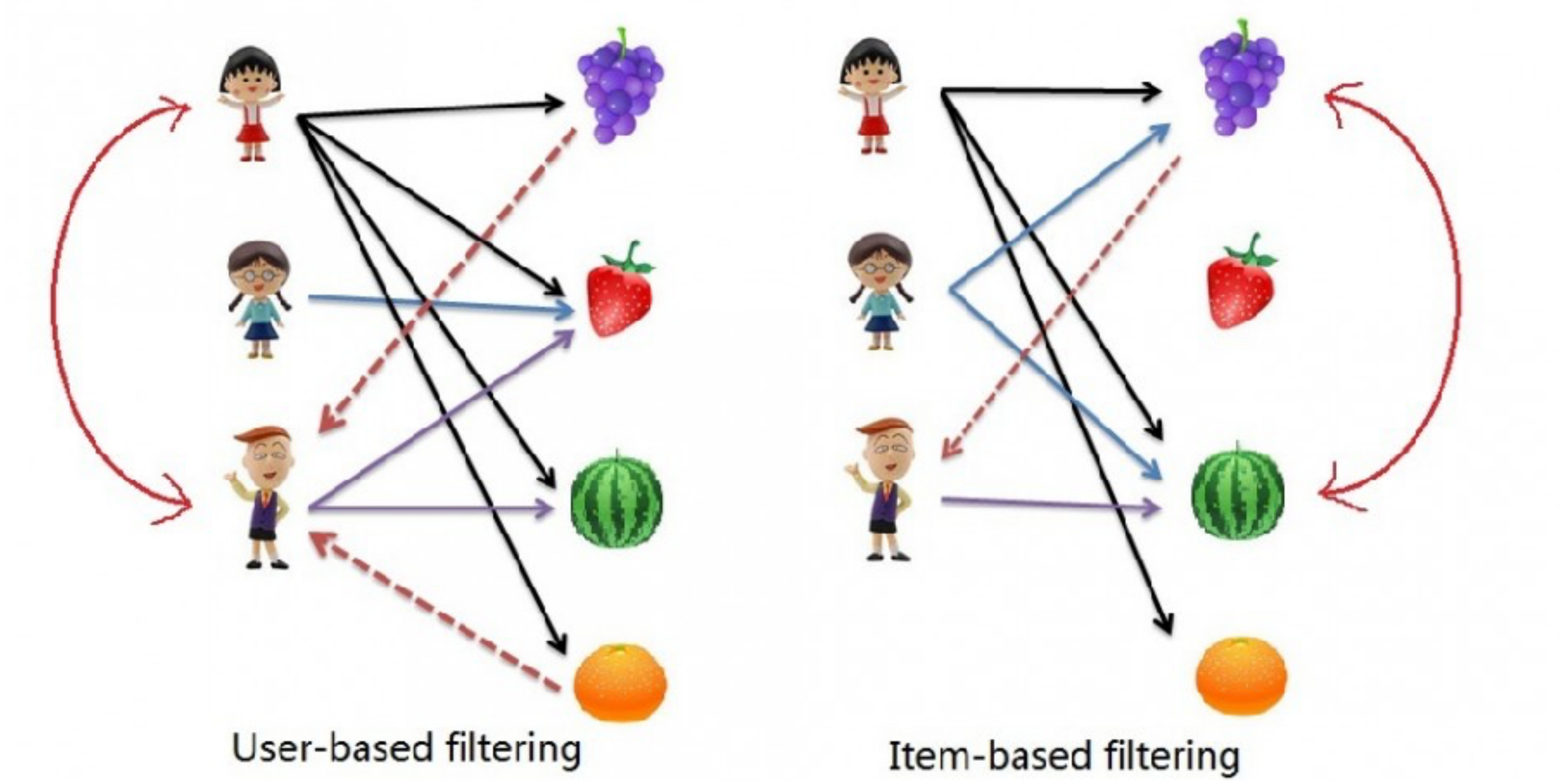
User-user:

find similar users to me and aggregate their ratings

Item-item:

find items that are similar the items I rated

Neighborhood-based



datalys

Model-based

Rule based

Latent factors

- **Matrix Factorization**
- **Neural Networks**

Matrix Factorization

Item

User	W	X	Y	Z
A		4.5	2.0	
B	4.0		3.5	
C		5.0		2.0
D		3.5	4.0	1.0

Rating Matrix

=

A	*1.2*	*0.8*
B	*1.4*	*0.9*
C	*1.5*	*1.0*
D	*1.2*	*0.8*

User Matrix

X

W	X	Y	Z
1.5	*1.2*	*1.0*	*0.8*
1.7	*0.6*	*1.1*	*0.4*

Item Matrix

Source:
http://jxieeducation.com/2016-06-26/Factorization-Machines-A-Theoretical-Introduction/

Group RS

Recommending items for a group of people such that everyone *should* be satisfied.

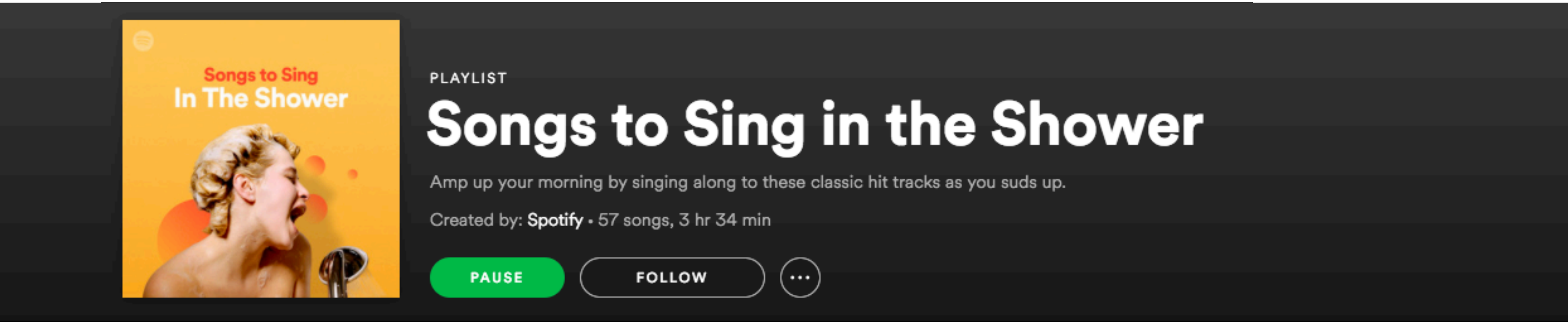
Domains:
travel, restaurants, movies.

How to *aggregate*?

datalys

Context-aware RS

context = conditions or circumstances which affect something



Contextual features:
time, location, purchasing purpose, mood, ...

datalys

Hybrid

A combination of different recommendation techniques.

EXAMPLE:

Recommend with *content-based* to a new user.

After receiving a specific amount of ratings (for this user), continue with *collaborative filtering*.

Several combination strategies.

Resources

IR @ fiit

Conferences: RecSys, UMAP

Useful datasets:

- Grouplens (https://grouplens.org/datasets/)
- YELP (https://www.yelp.com/dataset/challenge)

datalys

Sources

Ricci et al. Recommender Systems Handbook. 2015.