



# Recommender Systems

Personalization in Information Retrieval

# What is a Recommender System?

- ▶ Common definition -
  - ▶ A subclass of information retrieval systems that seek to predict the "*rating*" or "*preference*" that a user would give to an item.
- ▶ Idea: Use *items* as generic term for what is recommended.
  - ▶ Help people (customers, users) make *decisions*

# What is a Recommender System?

- ▶ Applications –
  - ▶ In eCommerce - recommend *products*
  - ▶ In eLearning/e-News - recommend *content*
  - ▶ In online social networks - recommend *friends, posts, jobs, music..*
  - ▶ In search and navigation - recommend *links, pages, services ...*

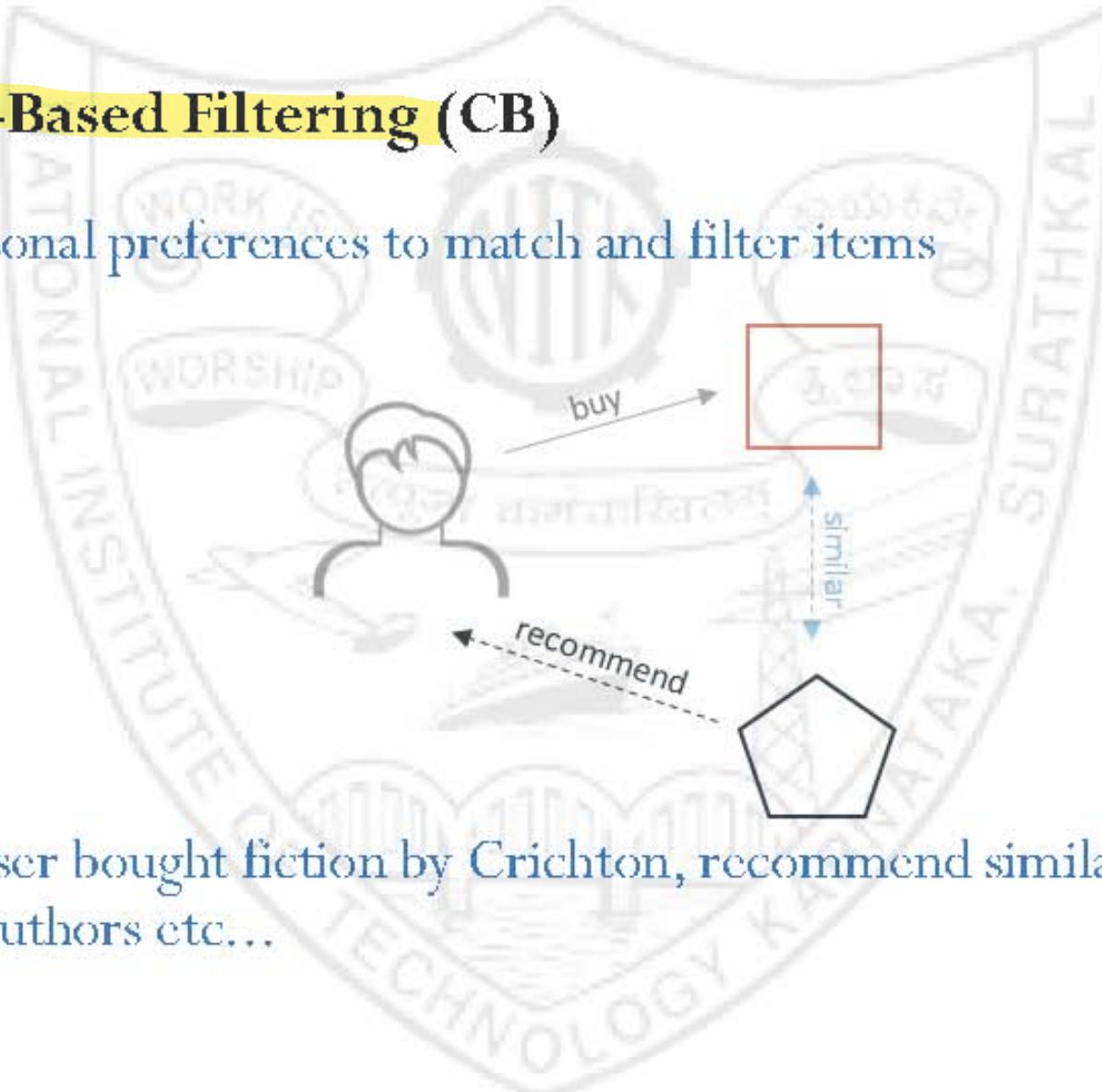
# Types of Recommenders

- ▶ Recommendation is based on preferences of -
  - ▶ An individual
  - ▶ A related group or community
  - ▶ An unrelated group or community

# Recommender Systems – Individual based

## Content-Based Filtering (CB)

- use personal preferences to match and filter items

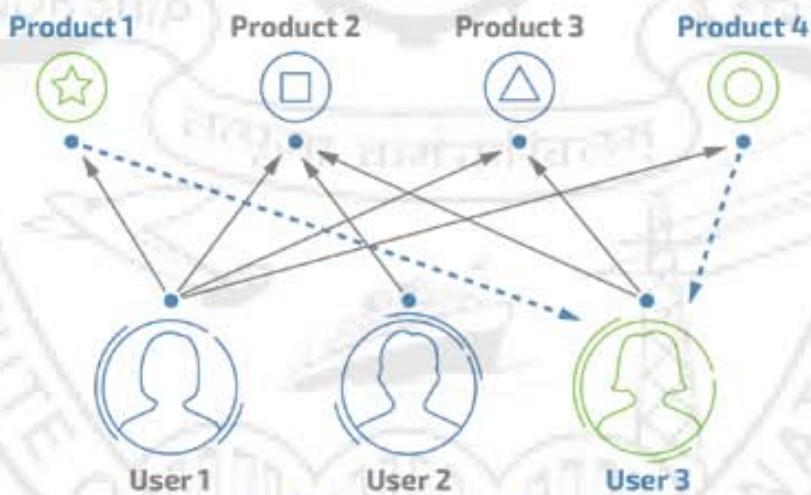


- E.g. A user bought fiction by Crichton, recommend similar books/ genre/authors etc...

# Recommender Systems – Community based

## ► Collaborative Filtering (CF)

- match 'like-minded' users.



- E.g. if two people have similar 'taste', recommend related items to them.

# Recommender Systems – Unrelated Group based

## ► Social Data Mining

- Mine log data of social activity to learn group preferences.



- E.g. web usage mining, for targeted advertising ...

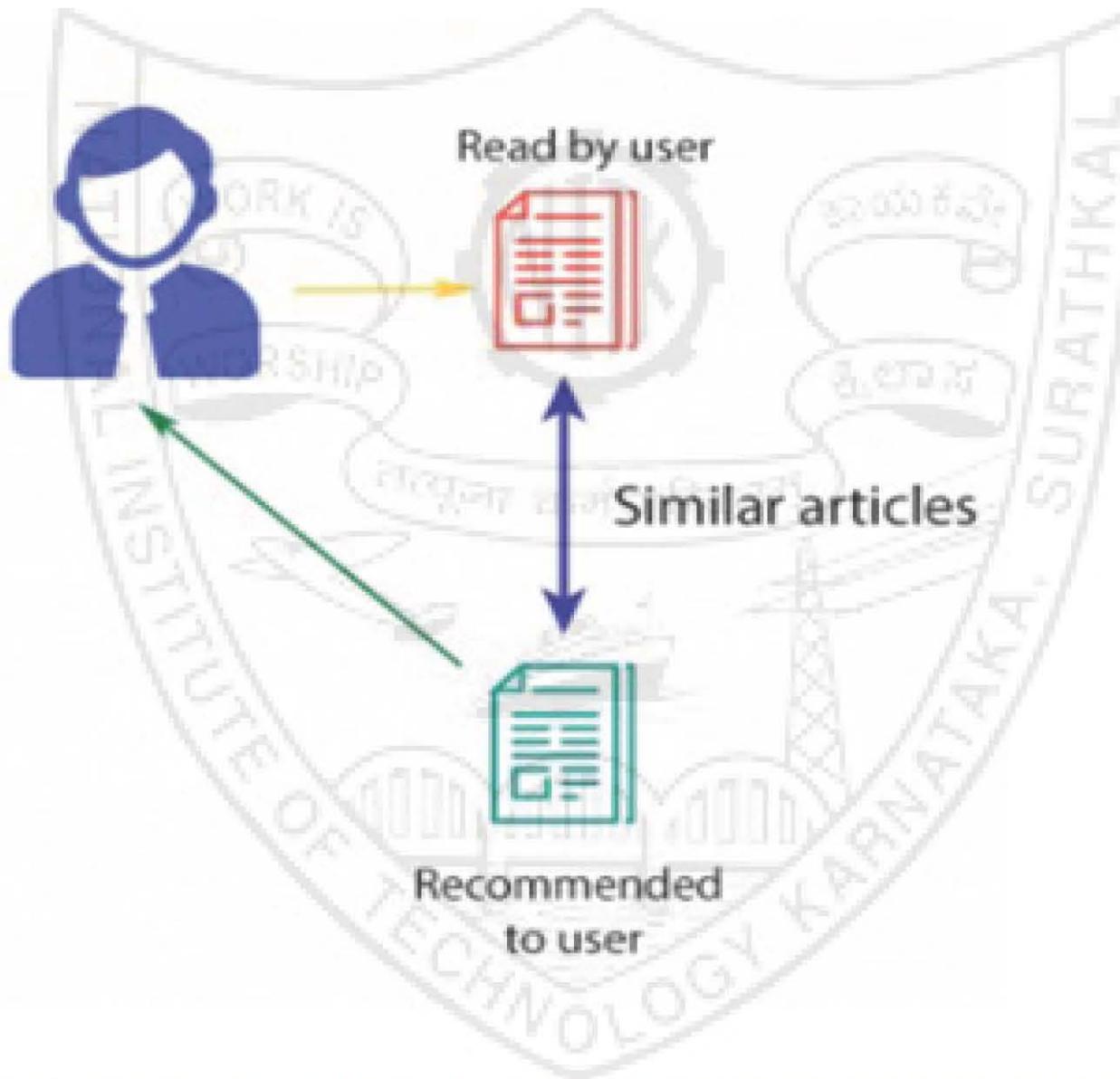
# Content-Based Recommenders



# Content-Based Recommenders (CB)

- ▶ Use ML to learn preferences through user feedback and build a user profile.
  - ▶ “Find me more/new things similar to the ones I have”
- ▶ Recommendation is viewed as a search process
  - ▶ *user profile* acting as the *query*
  - ▶ *set of items* acting as the *documents to match*.

# Content-Based Recommenders (CB)



# Content-Based Recommenders (CB)

- ▶ Methods adopted –
  - ▶ **Explicit feedback**
    - ▶ user actively rates items
  - ▶ **Implicit feedback – system records user activity**
    - ▶ Clickstream data classified according to page category and activity
      - e.g. browsing a product page
    - ▶ Time spent on an activity
      - E.g browsing a page

# Advantages of Content-Based Approach

- ▶ No cold-start problems.
- ▶ Able to recommend to users with unique tastes.
- ▶ Allows to address first rater bias
- ▶ Can provide explanations of recommended items by listing content-features that caused an item to be recommended.

# Disadvantages of Content-Based Method

- ▶ Requires content that can be encoded as *meaningful* features.
  - ▶ Users' tastes must be modeled as a *learnable* function of these content features.
- ▶ Unable to exploit quality judgments of other users, when required.

# Collaborative Filtering (CF)



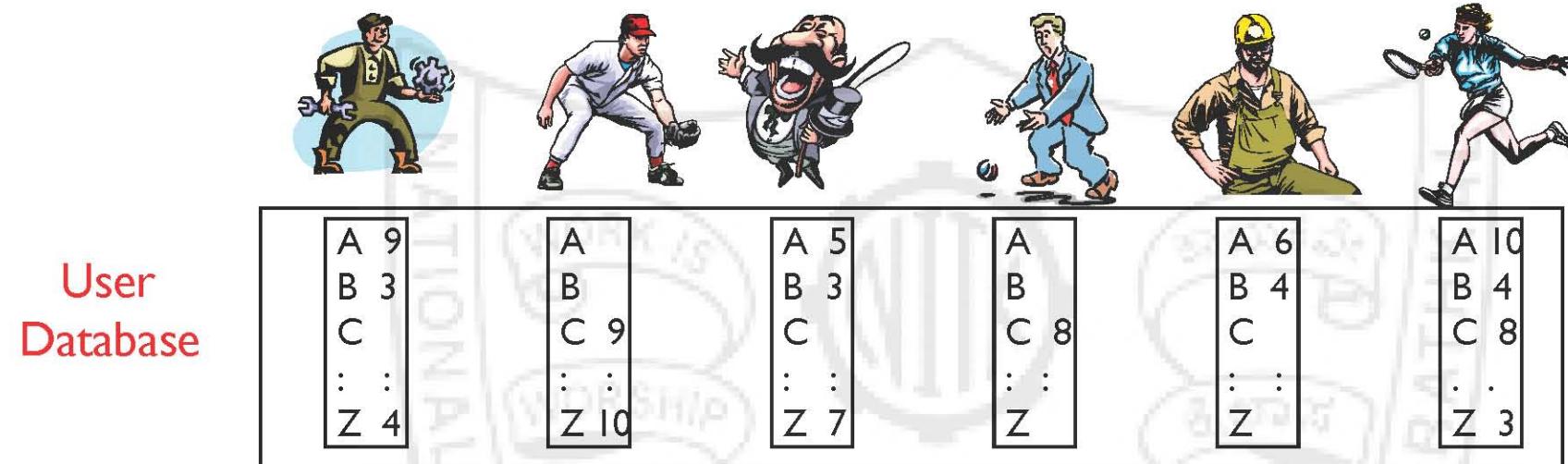
# Collaborative Filtering (CF)

- ▶ Methods adopted –
  - ▶ **Users rate items** – user interests recorded.
    - ▶ Ratings may be explicit or implicit.
  - ▶ **Nearest neighbour matching**
  - ▶ **Recommend highly rated items**
  - ▶ **Allow user to rate recommended items**
    - ▶ further enhances personalization.

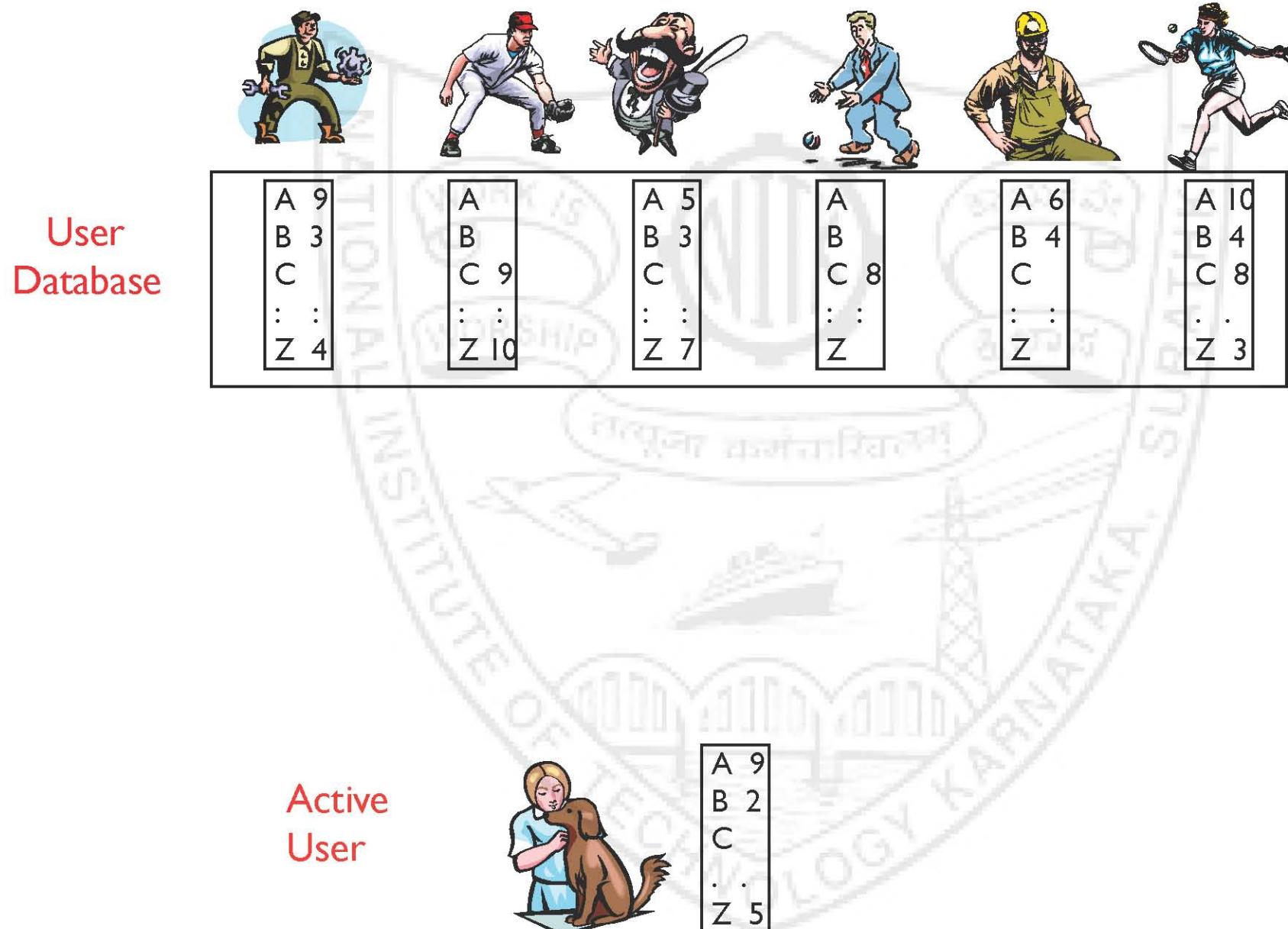
# Collaborative Filtering



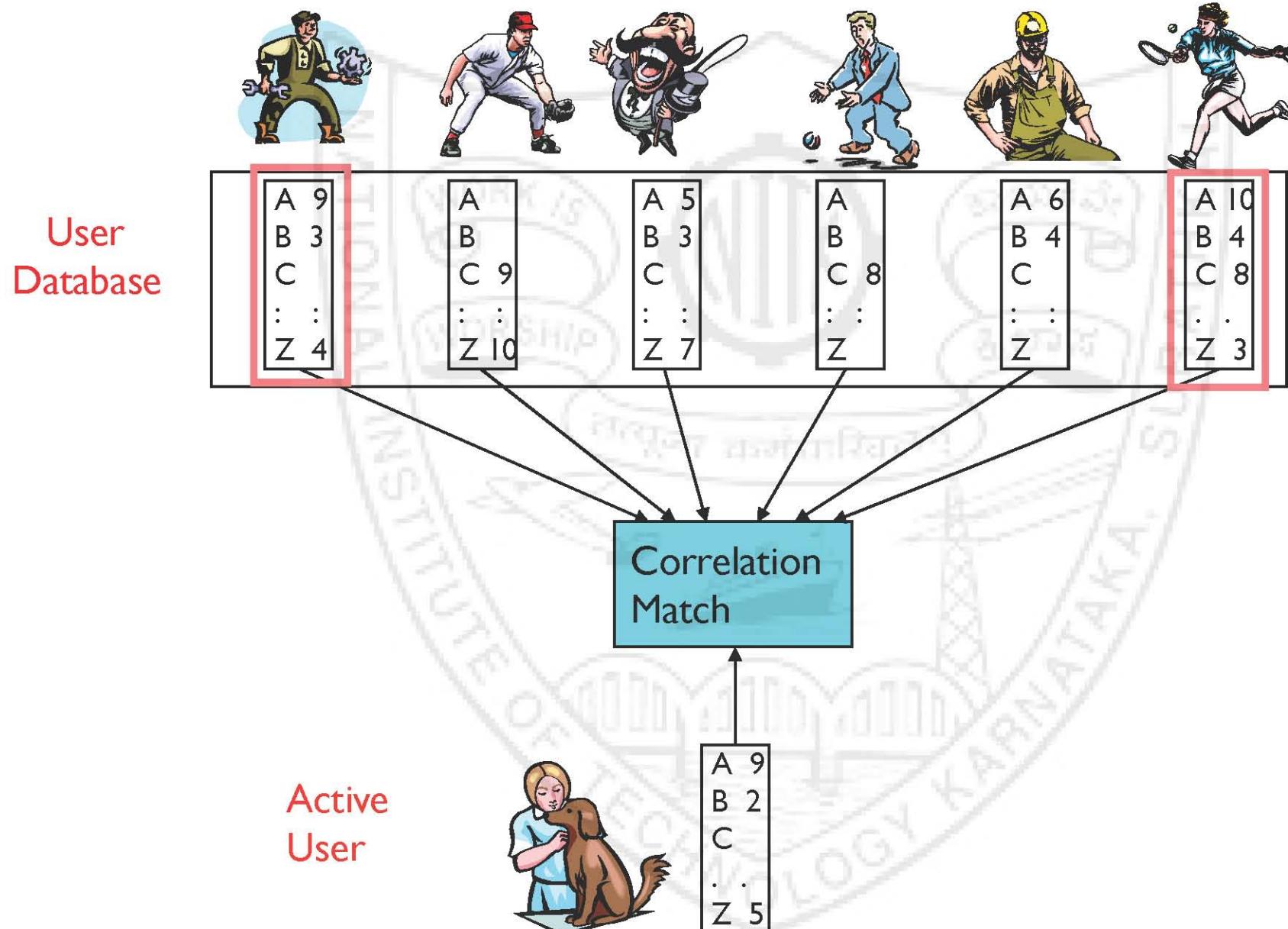
# Collaborative Filtering



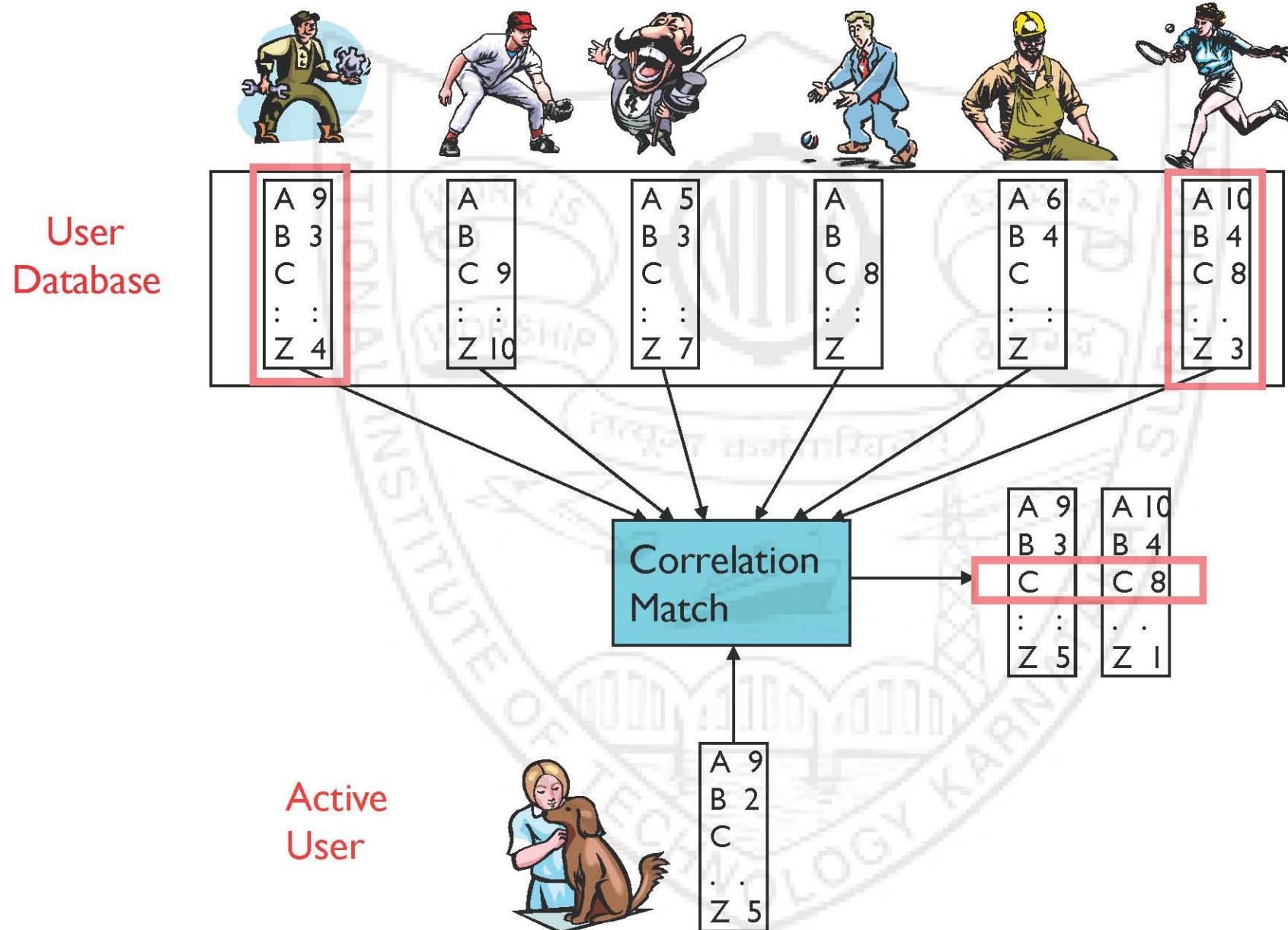
# Collaborative Filtering



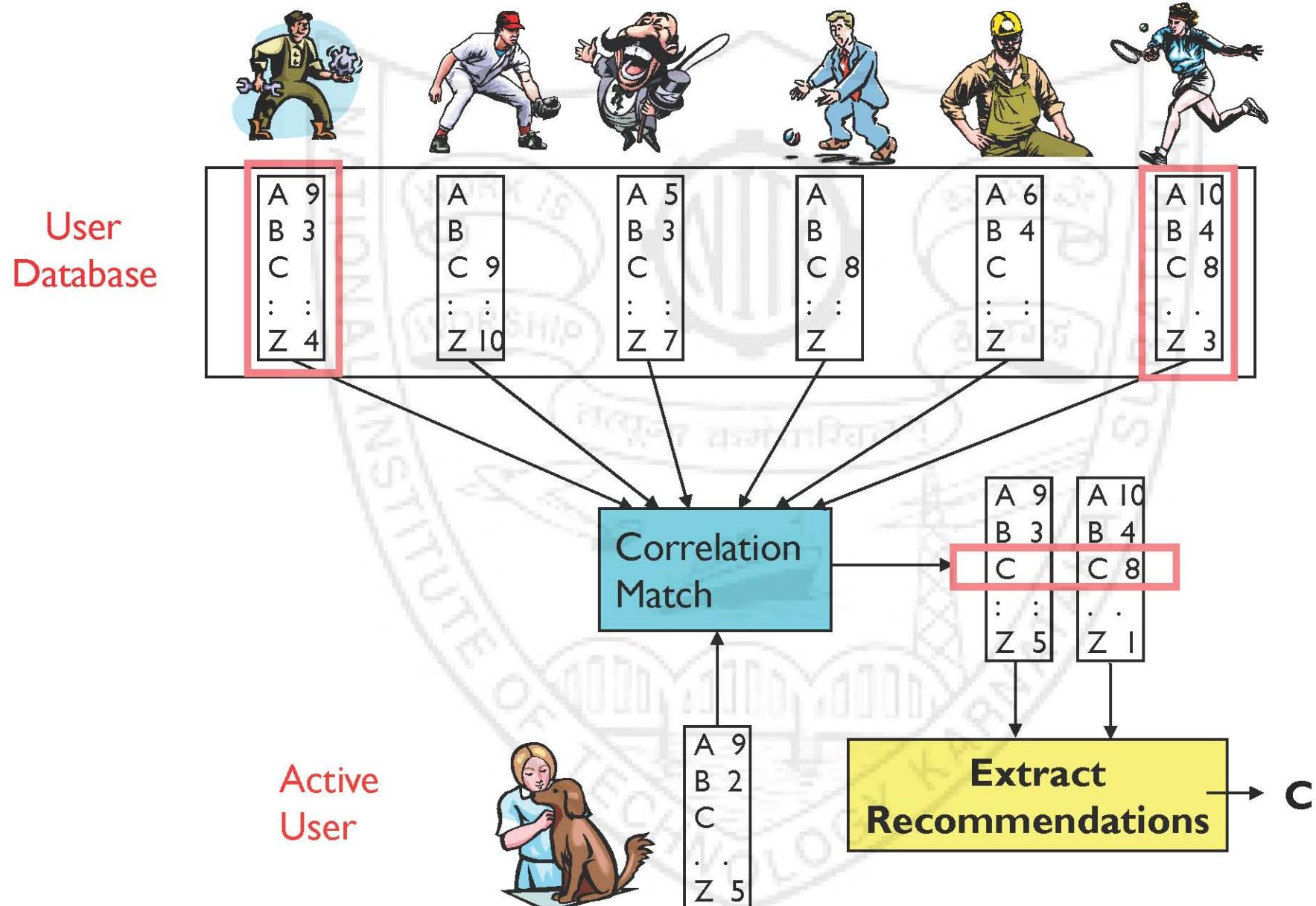
# Collaborative Filtering



# Collaborative Filtering



# Collaborative Filtering



# Similarity Weighting

- ▶ E.g. use Pearson correlation coefficient between ratings for active user,  $a$ , and another user,  $u$ .

$$C_{a,u} = \frac{\text{covar}(r_a, r_u)}{\sigma_{r_a} \sigma_{r_u}}$$

→  $r_a$  and  $r_u$  are the ratings vectors for the  $m$  items rated by both  $a$  and  $u$

→  $r_{i,j}$  is user  $i$ 's rating for item  $j$

# Covariance and Standard Deviation

## ► Covariance:

$$\text{covar}(r_a, r_u) = \frac{\sum_{i=1}^m (r_{a,i} - \bar{r}_a)(r_{u,i} - \bar{r}_u)}{m}$$

where,  $\bar{r}_x = \frac{\sum_{i=1}^m r_{x,i}}{m}$

## ► Standard Deviation:

$$\sigma_{r_x} = \sqrt{\frac{\sum_{i=1}^m (r_{x,i} - \bar{r}_x)^2}{m}}$$

# Neighbor Selection

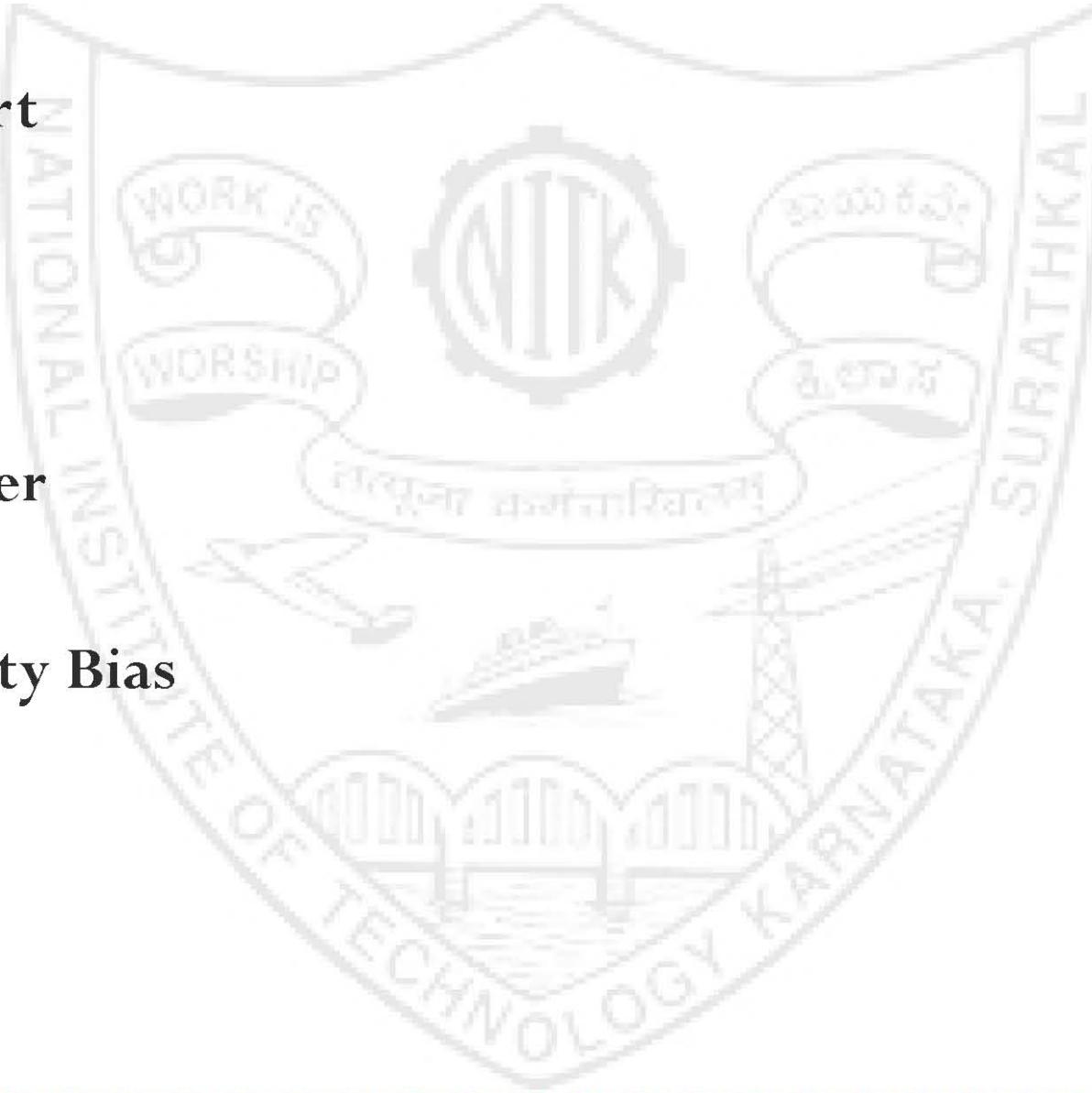
- ▶ For a given active user,  $a$ ,
  - ▶ select correlated users to serve as source of predictions.
  - ▶ Standard approach:
    - ▶ use the most similar  $n$  users,  $u$ , based on similarity weights,  $w_{a,u}$
  - ▶ Alternate approach:
    - ▶ include all users whose similarity weight is above a given threshold.

# Rating Prediction

- ▶ Predict a rating,  $p_{a,i}$ , for each item  $i$ , for active user,  $a$ , by using the  $n$  selected neighbor users,  $u \in \{1, 2, \dots, n\}$ .
- ▶ To account for users different ratings levels,
  - ▶ predictions are based on *differences* from a user's *average* rating.
- ▶ Weight users' ratings contribution by their similarity to the active user.

# Problems with Collaborative Filtering

- ▶ Cold Start
- ▶ Sparsity
- ▶ First Rater
- ▶ Popularity Bias



# Combining Content and Collaboration

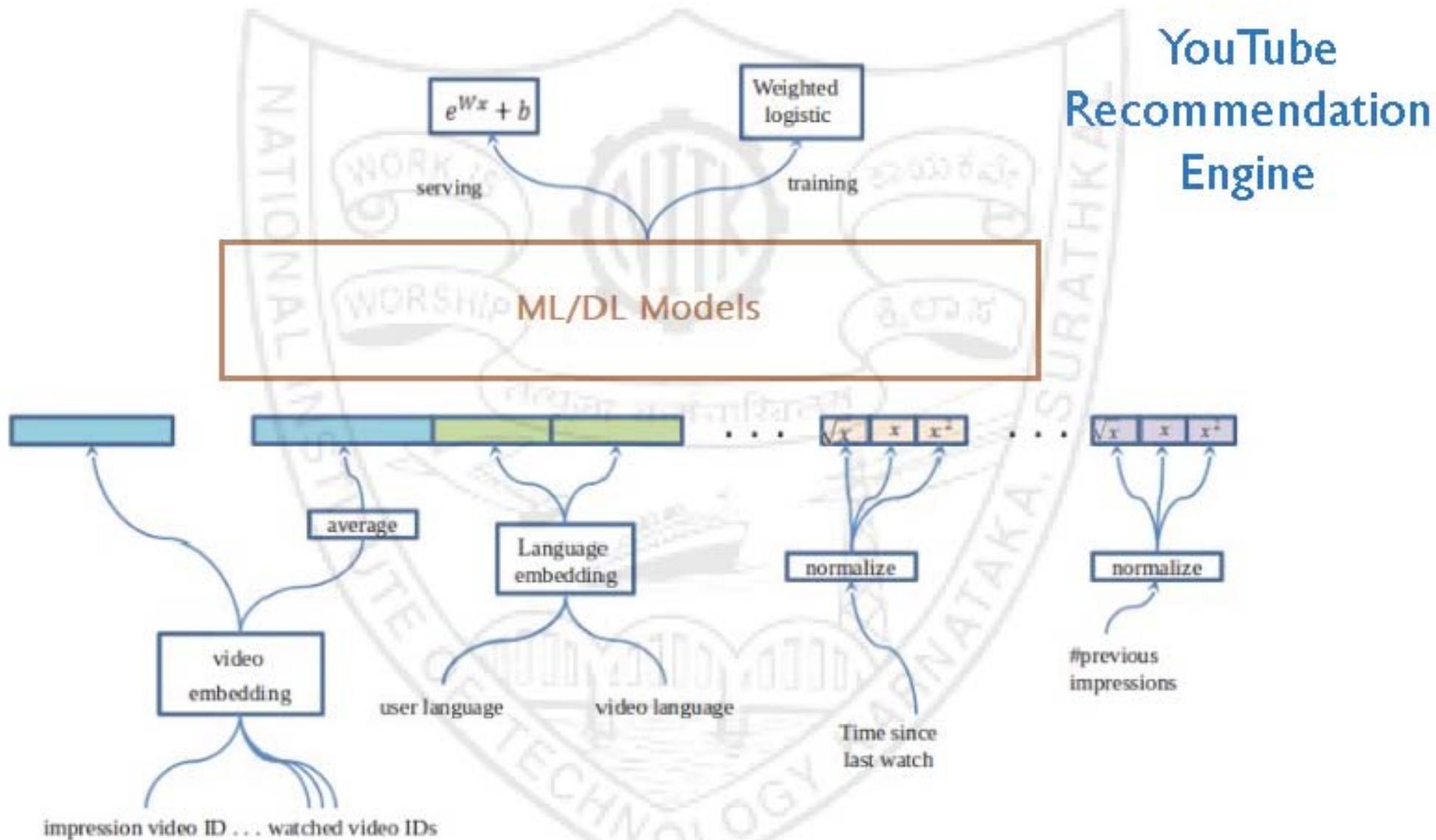
---

- ▶ Content-based and collaborative methods have complementary strengths and weaknesses.
- ▶ *Solution:* Combine methods to obtain the best of both.

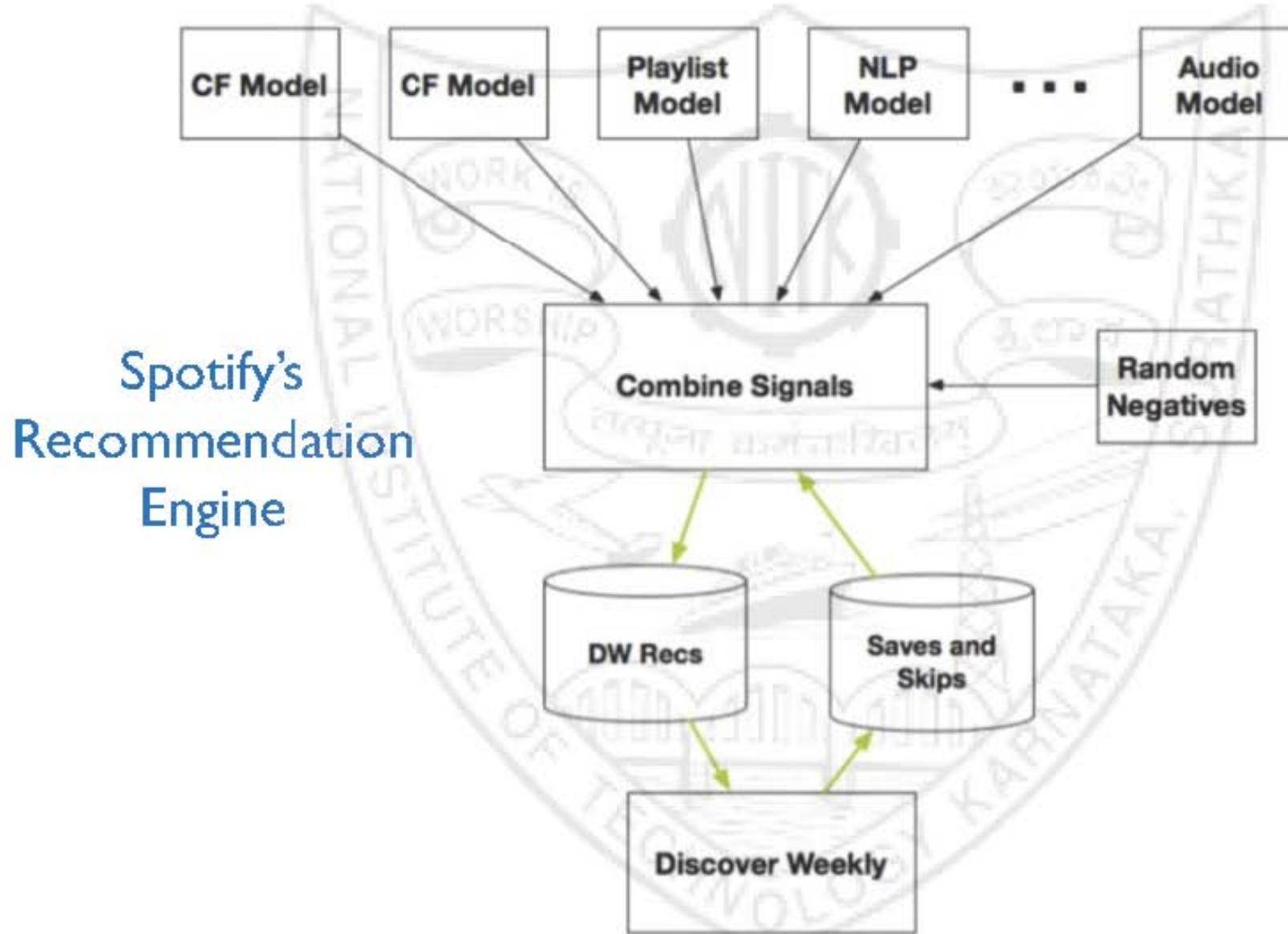
# Combining Content and Collaboration

- ▶ Various hybrid approaches:
  - ▶ Apply both methods and combine recommendations.
  - ▶ Use collaborative data as content.
  - ▶ Use content-based predictor as another collaborator.
  - ▶ Use content-based predictor to complete collaborative data.
  - ▶ ....

# Media-specific recommendation systems



# Media-specific recommendation systems



# Recommenders - Examples

## ► Amazon's Collaborative Filtering Algorithm The pioneer

### Frequently Bought Together



Price For All Three: \$258.02

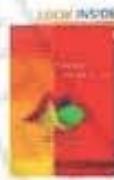
[Add all three to Cart](#)

- [This item: The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition \(Springer Series in Statistics\)](#) by Trevor Hastie
- [Pattern Recognition and Machine Learning \(Information Science and Statistics\)](#) by Christopher M. Bishop
- [Pattern Classification \(2nd Edition\)](#) by Richard O. Duda

### Customers Who Bought This Item Also Bought



[All of Statistics: A Concise Course in Statist...](#) by Larry Wasserman  
 (8) \$60.00



[Pattern Classification \(2nd Edition\)](#) by Richard O. Duda  
 (27) \$117.25



[Data Mining: Practical Machine Learning Tools an...](#) by Ian H. Witten  
 (29) \$41.55



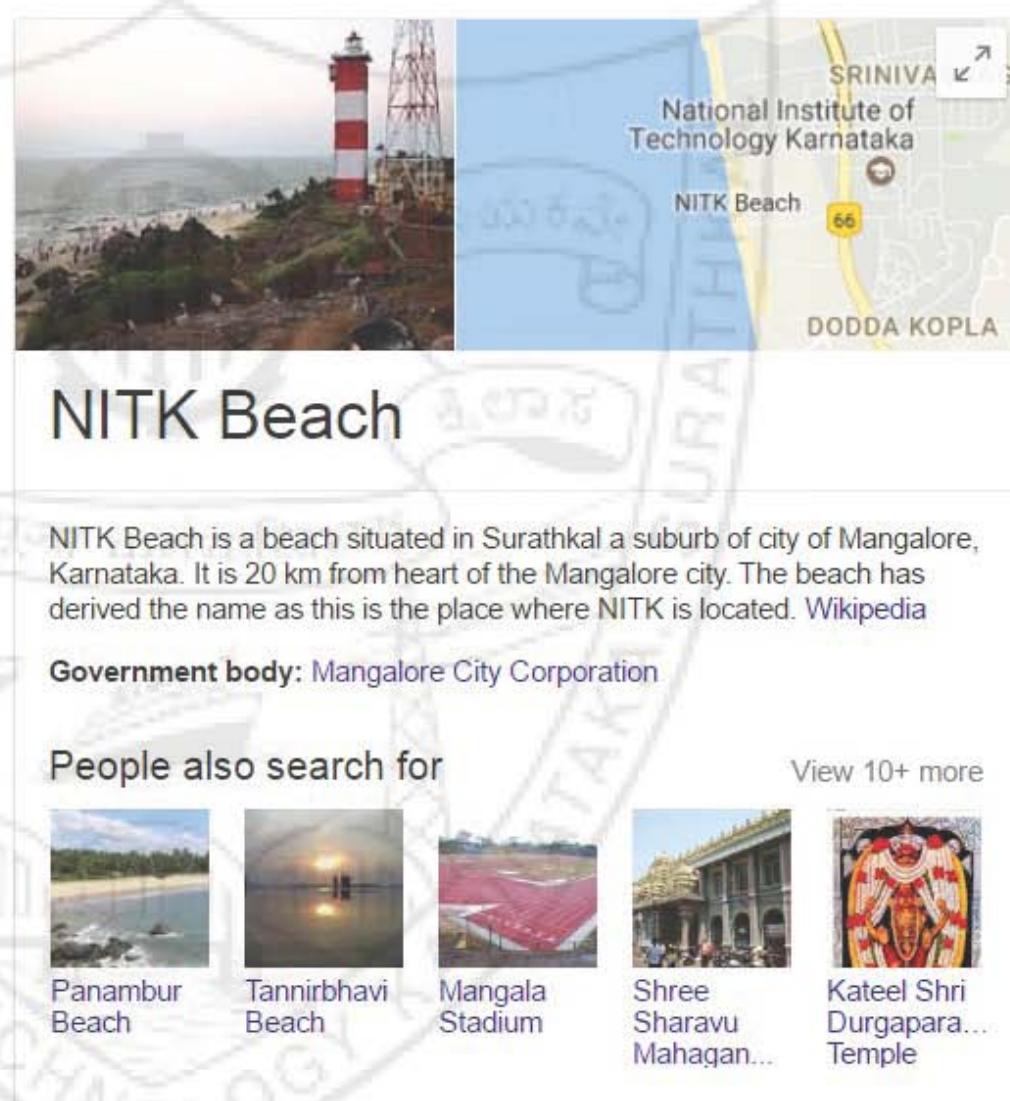
[Bayesian Data Analysis, Second Edition \(Texts in...](#) by Andrew Gelman  
 (10) \$56.20



[Data Analysis Using Regression and Multilevel /...](#) by Andrew Gelman  
 (13) \$39.59

# Recommenders - Some Examples

- ▶ Google Knowledge Graph



A screenshot of a Google Knowledge Graph card for "NITK Beach". The card features a large image of a red and white striped lighthouse on a hillside overlooking a beach. Below the image, the text "NITK Beach" is displayed in a large, bold, dark font. To the right of the text is a map showing the location of NITK Beach near the National Institute of Technology Karnataka (NITK). The map includes labels for "SRINIVAS", "National Institute of Technology Karnataka", "NITK Beach", and "DODDA KOPLA". A yellow route marker indicates a path from "SRINIVAS" to "NITK Beach". Below the title, a detailed description of NITK Beach is provided: "NITK Beach is a beach situated in Surathkal a suburb of city of Mangalore, Karnataka. It is 20 km from heart of the Mangalore city. The beach has derived the name as this is the place where NITK is located. [Wikipedia](#)". Underneath the description, it says "Government body: [Mangalore City Corporation](#)". A section titled "People also search for" lists several related locations with their names and small thumbnail images: Panambur Beach, Tannirbhavi Beach, Mangala Stadium, Shree Sharavu Mahagan..., and Kateel Shri Durgapara... Temple. There is also a link "View 10+ more". At the bottom right of the card is a "Feedback" button.

NITK Beach

NITK Beach is a beach situated in Surathkal a suburb of city of Mangalore, Karnataka. It is 20 km from heart of the Mangalore city. The beach has derived the name as this is the place where NITK is located. [Wikipedia](#)

Government body: [Mangalore City Corporation](#)

People also search for

Panambur Beach

Tannirbhavi Beach

Mangala Stadium

Shree Sharavu Mahagan...

Kateel Shri Durgapara... Temple

View 10+ more

Feedback

# Recommenders - Some Examples

Google movies similar to avatar

All News Images Videos Shopping More Tools Services

What to watch

Films Avatar Clear all

If you like Avatar: Creating the World of Pandora

War for the Plan... Mortal Engines Ready Player One Alita: Battle Angel Rise Of The Plan... Valerian and the ... Oblivion Rampage

Action films

About 25,20,00,000 results (1.18 seconds)

Feedback

The image shows a Google search results page for the query "movies similar to avatar". The search bar at the top contains the query. Below it, the standard Google navigation links (All, News, Images, Videos, Shopping, More, Tools, Services) are visible. A large, semi-transparent watermark for "NITK SURATHKAL" is overlaid across the entire page. The main content area is titled "What to watch" and features a section titled "If you like Avatar: Creating the World of Pandora". It displays eight movie posters: War for the Plan..., Mortal Engines, Ready Player One, Alita: Battle Angel, Rise Of The Plan..., Valerian and the ... (partially visible), Oblivion, and Rampage. Below this, another section titled "Action films" shows several more movie posters, including Big Meets Bigger, Valerian, and other titles. At the bottom left, it says "About 25,20,00,000 results (1.18 seconds)". On the right side, there is a "Feedback" link.

# Recommenders - Some Examples

Google movies similar to Kantara

All Images Shopping News Videos More Tools

About 2,21,00,000 results (0.64 seconds)

**Kantara**  
CBFC: U/A 2022 · Thriller/Adventure · 2h 30m

Overview Showtimes Reviews Cast

**More like this**

- 8.9. 777 Charlie. Watchlist.
- 8.4. K.G.F: Chapter 2. Watch options.
- 8.4. Garuda Gamana Vrishabha Vahana. Watchlist.
- 8.2. K.G.F: Chapter 1. Watch options.
- 9.2. Ramayana: The Legend of Prince Rama. Watchlist.
- 8.5. Sita Ramam. Watch options.
- 7.4. Doctor G. Watch options.
- 7.9. Karthikeya 2. Watchlist.

More items...

<https://www.imdb.com/title/tt1111111/>

**Kantara (2022) - IMDb**

Get tickets

About

9.1/10 · IMDb

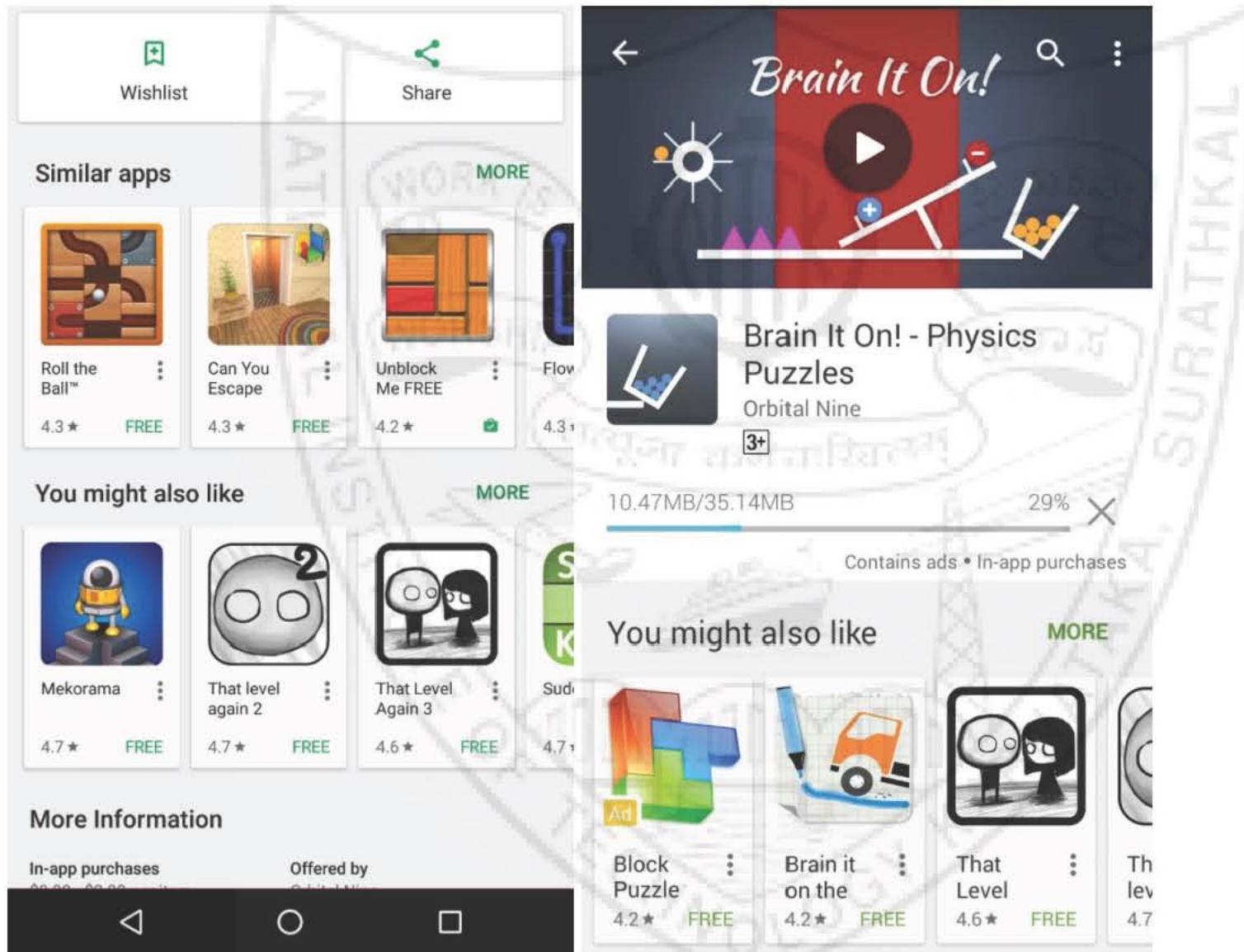
97% liked this film

Google users

Kantara is a 2022 Indian thriller film written and directed by Vittu. It stars Sharwanand, Nivetha Thomas, and Sharwanand's son Sharwanand. The film is produced by Vittu and is set in the backdrop of the 1921 Mysore Massacre. The story follows a group of revolutionaries who plan to assassinate the British ruler of Mysore. The film features a mix of action, drama, and romance. It has received critical acclaim for its direction, acting, and visual effects. The soundtrack, composed by Himesh Reshammiya, is also highly praised. The film has won several awards at various film festivals and has become a commercial success.

About featured snippets · Feedback

# Recommenders - Some Examples



Google Play

# Recommenders - Some Examples

- ▶ Facebook Open Graph enabled applications

The screenshot shows the Facebook search interface with a sidebar on the left and a main search results area on the right.

**Left Sidebar:**

- Search bar: Search for people, places and things
- People I may know
- My friends** (highlighted)
- Photos I might like
- Music I may like
- Movies I may like
- Nearby restaurants
- Games I may like
- Photos I have liked

**Center:**

Refine This Search

- Photos of: Add... ▾
- Location: Add... ▾
- Time: Add... ▾
- Commented on by: Yiyang Yang ▾ Add
- Liked by: Add... ▾
- Photos by: Add... ▾

SEE MORE FILTERS

**Right Side:**

Yiyang Yang

**Profile:** Yiyang Yang  
Tallinn, Estonia

Photos of Yiyang Yang

Photos Yiyang Yang likes

Friends of Yiyang Yang

Pages Yiyang Yang likes

Photos Yiyang Yang commented on

Find all people named "Yiyang Yang"

See more results for "Yiyang Yang"

# Recommenders - Some Examples

- ▶ Facebook Open Graph enabled applications



# Recommenders - Some Examples

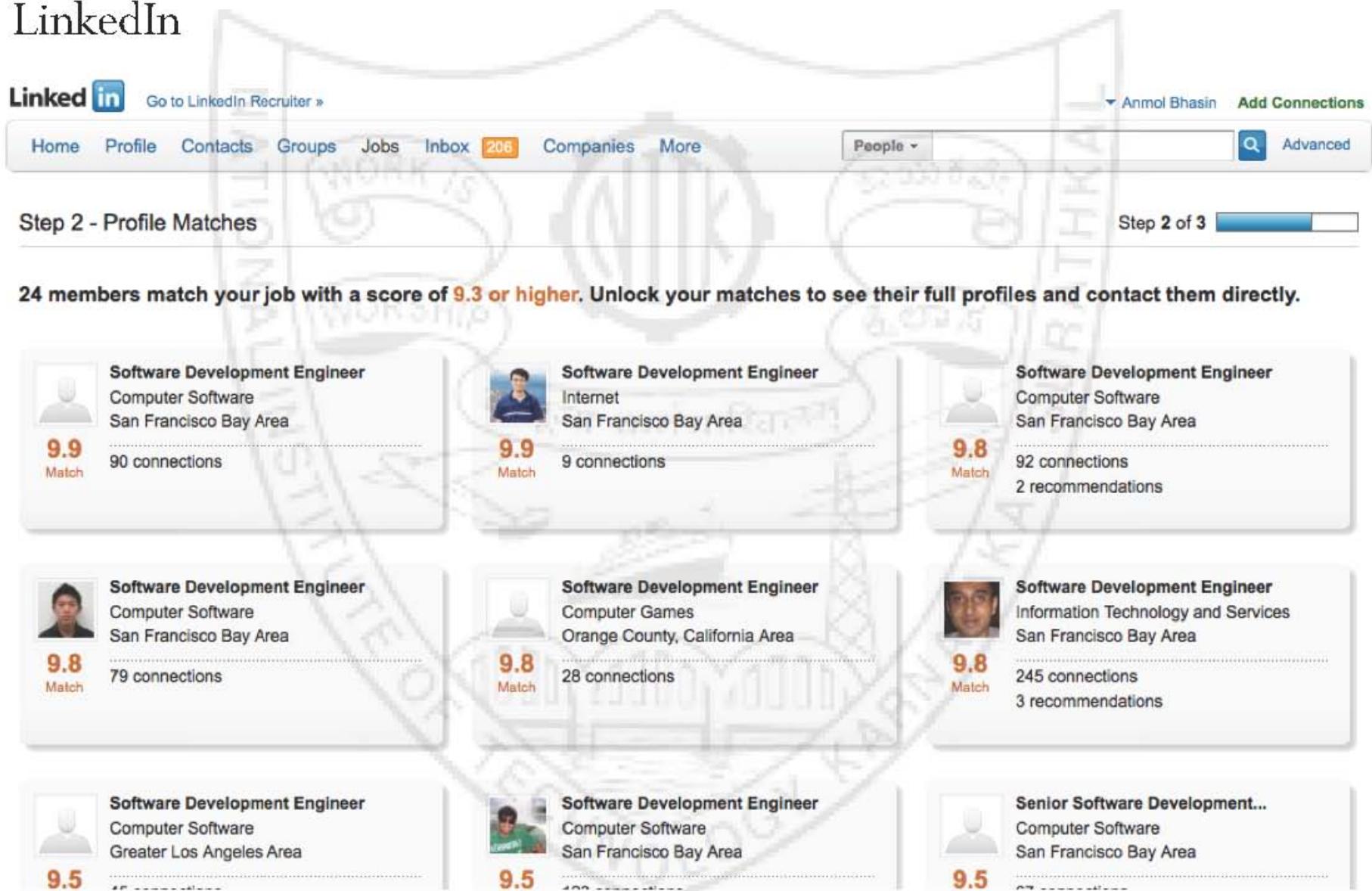
## ► Twitter

The image shows a screenshot of the Twitter mobile application's 'Getting started' guide, specifically Step 3 of 5. The header reads 'Getting started' and 'Step 3 of 5'. The main content area is titled 'Suggestions just for you.' and includes the text: 'Based on your choices, here are some suggestions for you. We recommend [following](#) all of them!'. Below this, there is a section titled 'Suggestions for you' with a list of tweets:

- John Gruber @gruber** · Aug 28  
"She's like a drunk older woman on a cruise."  
📍 The Las Vegas Strip, Paradise
- Walt Mossberg @waltmossberg** · Aug 29  
Apple Built iOS 8, Developers Are Coming [on.recode.net/1rylLYI](http://on.recode.net/1rylLYI) via [@InaFried](#) & [@DawnC331](#)
- Danny Sullivan @dannysullivan** · 20h  
I can confirm saying "Xbox load Battlefield no wait scratch that load Titanfall" works. I found that pretty impressive.
- Om Malik @om** · 1h  
Because I am missing Paris and my Parisian friends! [#latergram](#) [#endofsummer](#) [instagram.com/p/sVVRUUqThF/](http://instagram.com/p/sVVRUUqThF/)
- Jenna Wortham Δ @jennydeluxe** · 24h  
Some thoughts on Instagram's Hyperlapse, and Fast-Forwarding to the Future [nyti.ms/1pqQfcZ](http://nyti.ms/1pqQfcZ)
- dick costolo @dickc** · Aug 26  
Snapchat at \$10b not absurd. Crazy growth, clear monetization path, & one of the best social product thinkers out there. Long (figuratively)
- Laughing Squid @LaughingSquid** · 1h  
Iris the Tiny Piglet Bounds Through the Grass, Gets a Belly Rub and Goes for a Walk Around the Neighborhood [laughingsquid.com/iris-the-tiny-...](http://laughingsquid.com/iris-the-tiny-...)
- Cult of Mac @cultofmac** · 22m

# Recommenders - Some Examples

## ▶ LinkedIn



The screenshot shows a LinkedIn search interface for 'Profile Matches'. At the top, there's a navigation bar with 'Home', 'Profile', 'Contacts', 'Groups', 'Jobs', 'Inbox' (with 206 notifications), 'Companies', and 'More'. A search bar includes 'People' and 'Advanced' filters. The main content area is titled 'Step 2 - Profile Matches' and shows 'Step 2 of 3'. It displays 24 profile matches for a job search. Each match card includes a profile picture, name, job title, location, connection count, and recommendation count. The matches are arranged in a grid.

Profile Picture	Name	Job Title	Location	Connections	Recommendations
	Software Development Engineer	Computer Software	San Francisco Bay Area	90 connections	Match
	Software Development Engineer	Internet	San Francisco Bay Area	9 connections	Match
	Software Development Engineer	Computer Software	San Francisco Bay Area	92 connections	2 recommendations
	Software Development Engineer	Computer Software	San Francisco Bay Area	79 connections	Match
	Software Development Engineer	Computer Games	Orange County, California Area	28 connections	Match
	Software Development Engineer	Information Technology and Services	San Francisco Bay Area	245 connections	3 recommendations
	Software Development Engineer	Computer Software	Greater Los Angeles Area	Match	
	Software Development Engineer	Computer Software	San Francisco Bay Area	100 connections	Match
	Senior Software Development...	Computer Software	San Francisco Bay Area	Match	

# Recommenders - Some Examples

## ▶ Last.fm

The screenshot shows the Last.fm Recommendation Dashboard. At the top, there's a navigation bar with links for Music, Users, Listen, Events, Widgets (highlighted in blue), Download, and a user profile section. Below the navigation is a search bar and a "Music Search" button.

The main content area is titled "My Recommendation Dashboard". It features a "Latest Recommendation" section where "cpt\_blasphemy" recommends the track "My Way" by Ben Folds. A message says, "hehehehehehehehe I love this version!" and it was posted 29 days ago. To the right, a welcome message reads: "Welcome, smange. This is your personal dashboard. We've stocked it full of real-time features and recommendations you might enjoy."

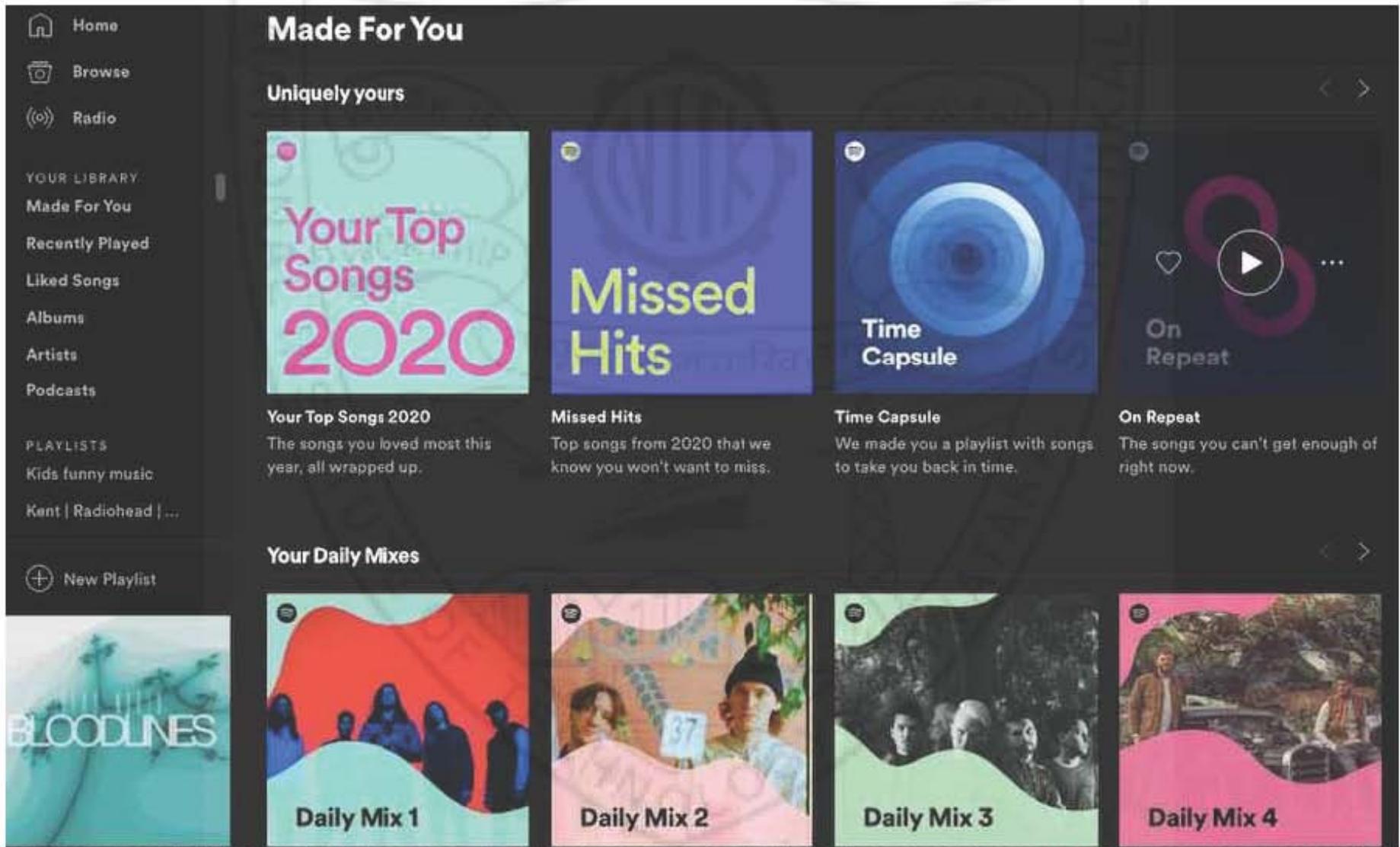
On the left, there's a "Shortcuts" sidebar with links to "My profile page", "Start Last.fm radio", "Find my friends", and "My settings". Below that is a "New To My Profile" sidebar showing recent activity: "Last Scrubbed Computer Games by Endorphin", "Most Recent Shout" from "cpt\_blasphemy" about Ebenezer leaving work, "Most Recent Visitor" from "cpt\_blasphemy" 6 hours ago, and "Recently Loved Goodbye".

The right side of the dashboard displays a "My Friends Online" section with a grid of friend profiles and their current listening tracks. A table lists these friends along with the tracks they are listening to and the time since they were last seen:

Friend	Track	Last Seen
thomas	Endorphin – Computer Games	just listened
harmz	Karen Ramirez – Troubled Girl (Spanish version)	just listened
thomae	Ben Folds – Gracie	2 minutes ago
harmz	Afterlife – 5th & Avenida	5 minutes ago
AlexCoombe	Everything but the Girl – I Don't Want to Talk About It	just listened
	Everything but the Girl – One Place	11 minutes ago
bjnice	VAST – Free	6h and 25m ago
	VAST – The Last One Alive	6h and 28m ago
Winterrmuzt	Skinny Puppy – Rivers	11h and 38m ago
	Skinny Puppy – Tin Omen	11h and 42m ago
TelekSterling	Elvis Presley – Can't Help Falling In Love	1d, 4h and 40m ago

# Recommenders - Some Examples

## ▶ Spotify



# Further Reading

---

- ▶ Resnick, Paul, and Hal R. Varian. "Recommender systems." *Communications of the ACM* 40.3 (1997): 56-58.
- ▶ Lops, Pasquale, Marco De Gemmis, and Giovanni Semeraro. "Content-based recommender systems: State of the art and trends." *Recommender systems handbook*. Springer, Boston, MA, 2011. 73-105.
- ▶ Adomavicius, Gediminas, and YoungOk Kwon. "Multi-criteria recommender systems." *Recommender systems handbook*. Springer, Boston, MA, 2015. 847-880.