# Erabiltzaile-ereduak, moldagarritasuna eta gomendioak

## *User models, adaptation, and recommendation*

Ainhoa Alvarez, Ana Arruarte, Mikel Larrañaga
Lengoaia eta Sistema Informatikoak saila

eman ta zabal zazu
Universidad del País Vasco
Euskal Herriko Unibertsitatea

# Content-based recommenders

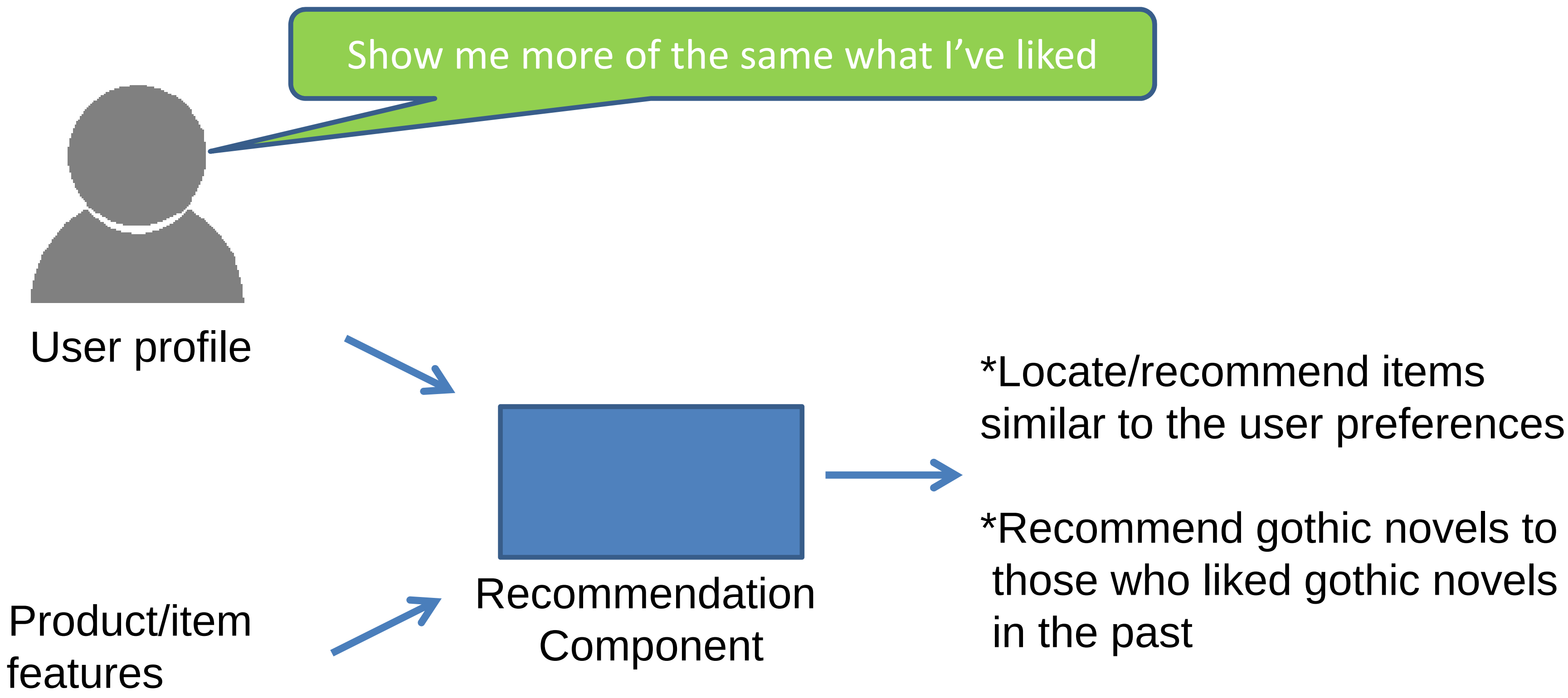
## *Edukian oinarritutako GS*

Ainhoa Alvarez, Ana Arruarte, Mikel Larrañaga
Lengoaia eta Sistema Informatikoak saila

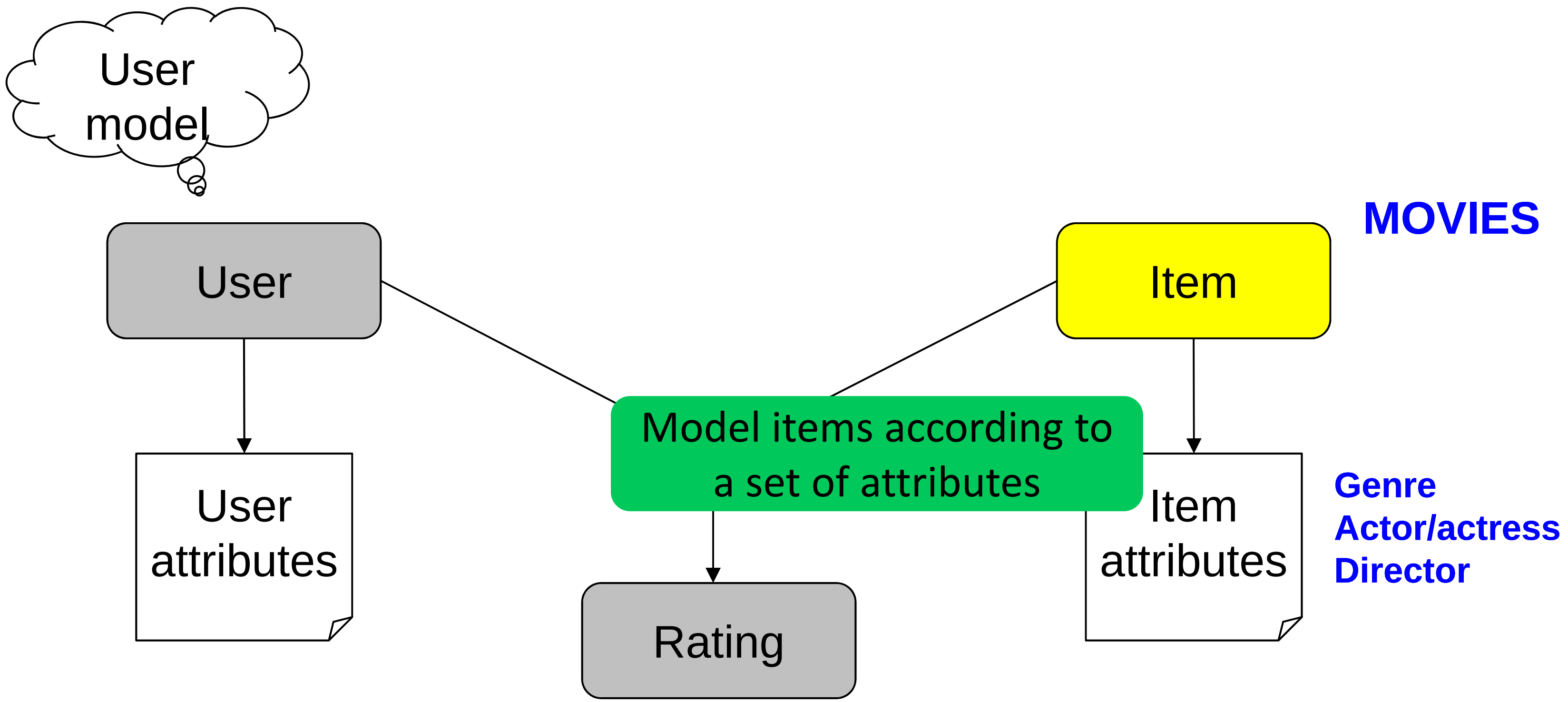
eman ta zabal zazu

Universidad del País Vasco
Euskal Herriko Unibertsitatea

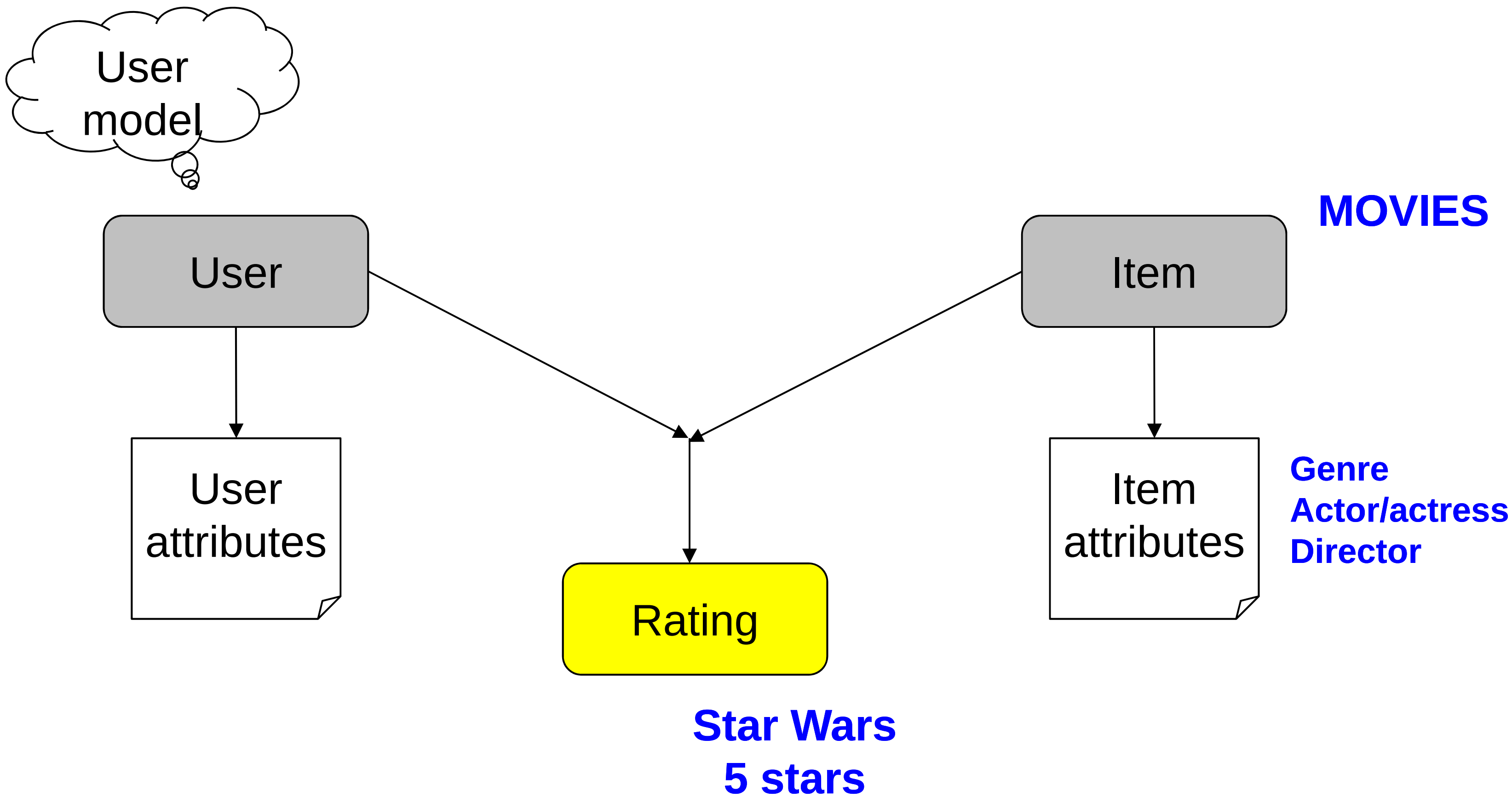
# Basic ideas


Show me more of the same what I’ve liked
User profile
Product/item features
Recommendation Component
*Locate/recommend items similar to the user preferences
*Recommend gothic novels to those who liked gothic novels in the past

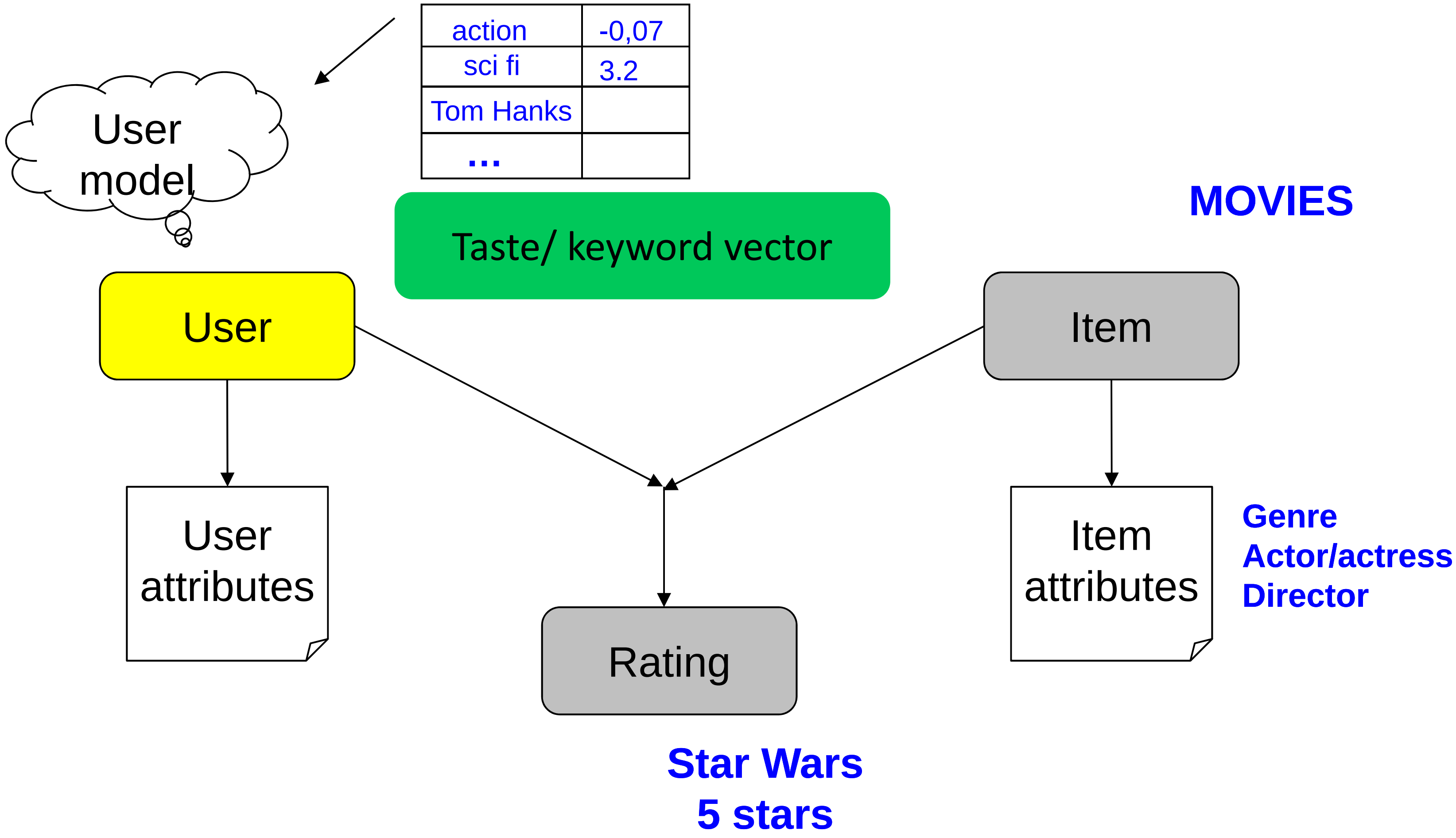
# Basic ideas, items



# Basic ideas, ratings



# Basic ideas, user model



# Basic ideas, resume

**Key concept:** Build a vector of atrribute preferences

Recommendation:
How much do I like a movie I have not seen before?
Mapping of the item attibutes and the user model

User profile

| | |
|---|---|
| action | -0,07 |
| sci fi | 3.2 |
| Tom Hanks | |
| … | |

Product/item features

| | |
|---|---|
| action | |
| sci fi | |
| Tom Hanks | |
| … | |

Recommendation Component

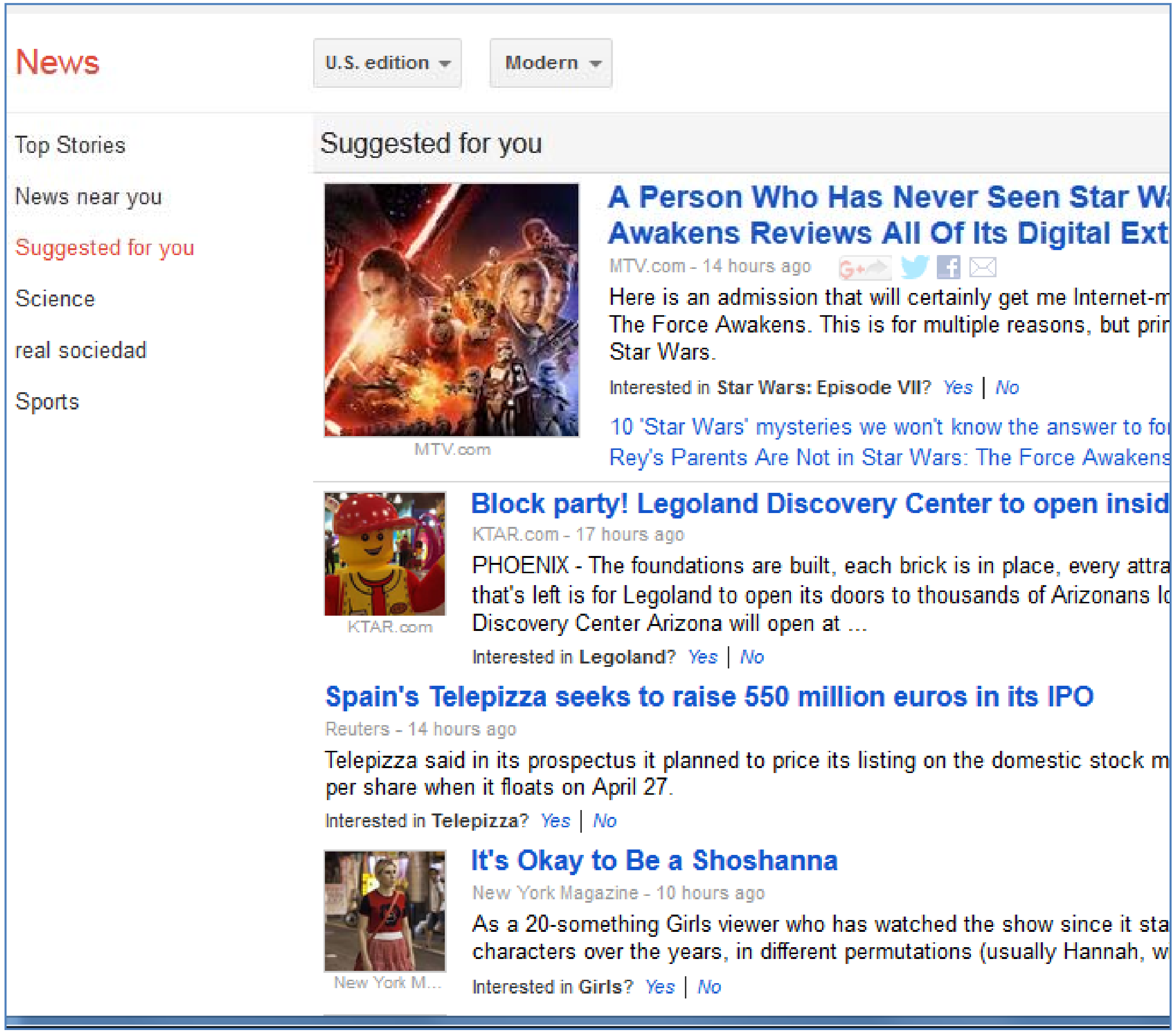
| item | score |
|---|---|
| i1 | |
| i2 | |
| … | |

Recommendation list

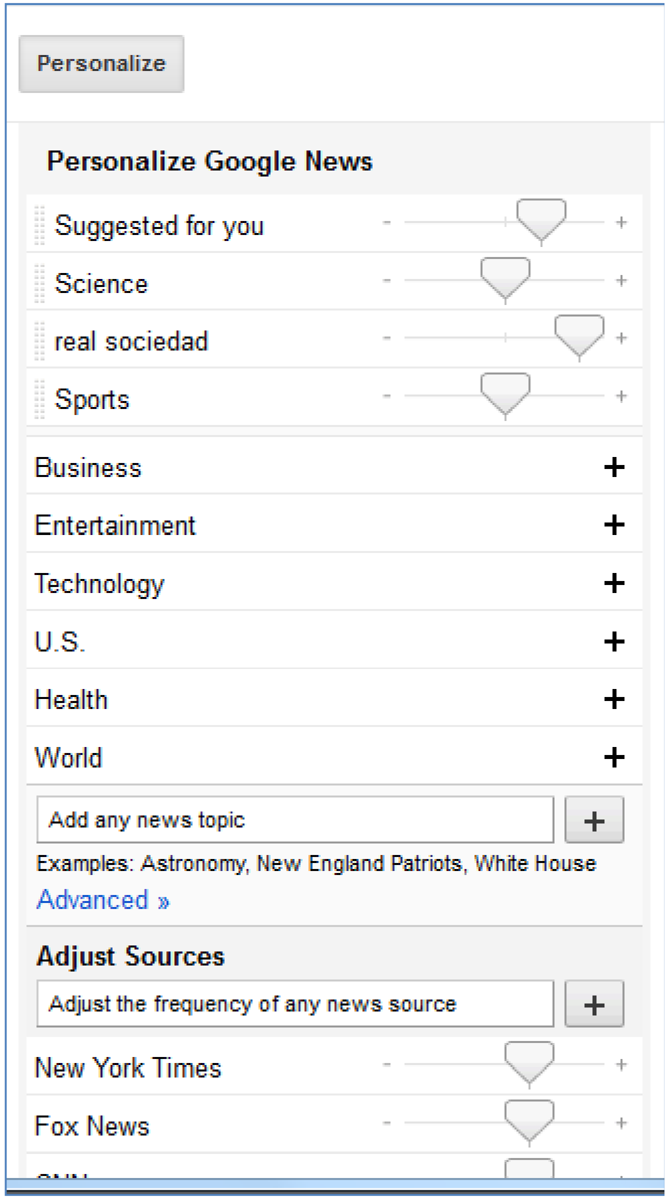
# Examples

- Personalized news feeds
- Artist of Genre music feeds

# Google News



News
U.S. edition
Modern
Top Stories
News near you
Suggested for you
Science
real sociedad
Sports
Suggested for you
A Person Who Has Never Seen Star W
Awakens Reviews All Of Its Digital Ext
MTV.com - 14 hours ago
Here is an admission that will certainly get me Internet-
The Force Awakens. This is for multiple reasons, but pri
Star Wars.
Interested in Star Wars: Episode VII? Yes | No
10 'Star Wars' mysteries we won't know the answer to fo
Rey's Parents Are Not in Star Wars: The Force Awakens
MTV.com
Block party! Legoland Discovery Center to open insid
KTAR.com - 17 hours ago
PHOENIX - The foundations are built, each brick is in place, every attra
that's left is for Legoland to open its doors to thousands of Arizonans l
Discovery Center Arizona will open at ...
Interested in Legoland? Yes | No
KTAR.com
Spain's Telepizza seeks to raise 550 million euros in its IPO
Reuters - 14 hours ago
Telepizza said in its prospectus it planned to price its listing on the domestic stock m
per share when it floats on April 27.
Interested in Telepizza? Yes | No
It's Okay to Be a Shoshanna
New York Magazine - 10 hours ago
As a 20-something Girls viewer who has watched the show since it sta
characters over the years, in different permutations (usually Hannah, w
Interested in Girls? Yes | No
New York M...

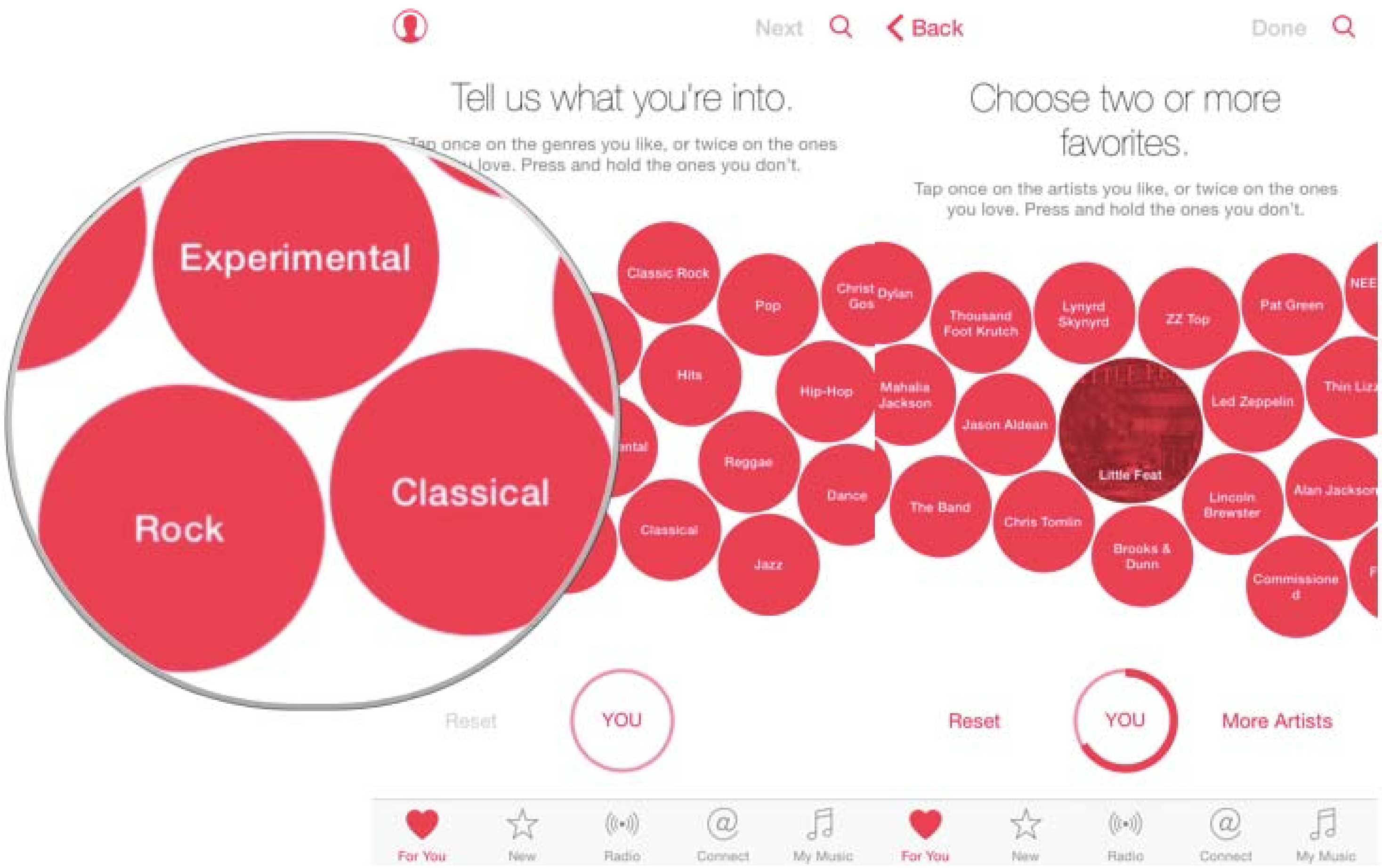



Personalize
Personalize Google News
Suggested for you
Science
real sociedad
Sports
Business
Entertainment
Technology
U.S.
Health
World
Add any news topic
Examples: Astronomy, New England Patriots, White House
Advanced »
Adjust Sources
Adjust the frequency of any news source
New York Times
Fox News

# Netflix


Just one more step to find movies you'll
Rate more so Netflix can get smarter about your movie preferences.

# “For You” section of Apple Music

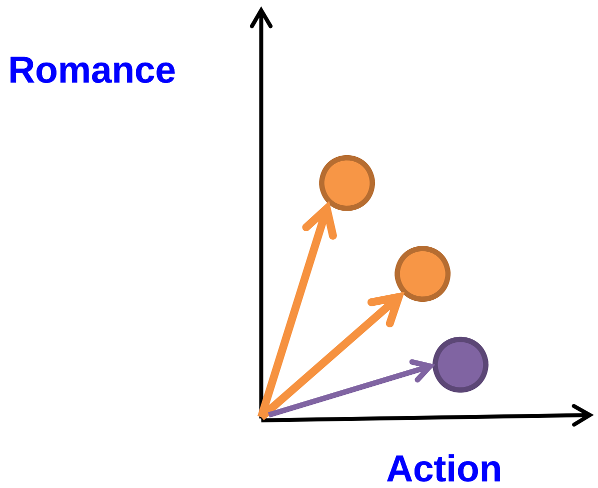


# Keyword vector

- Each keyword is a dimension
- Each item has a position in that space
- The position defines a vector
- Each user has a taste profile (also defines a vector)
- The match between user

preference and items is measured

by how closely the two vectors align

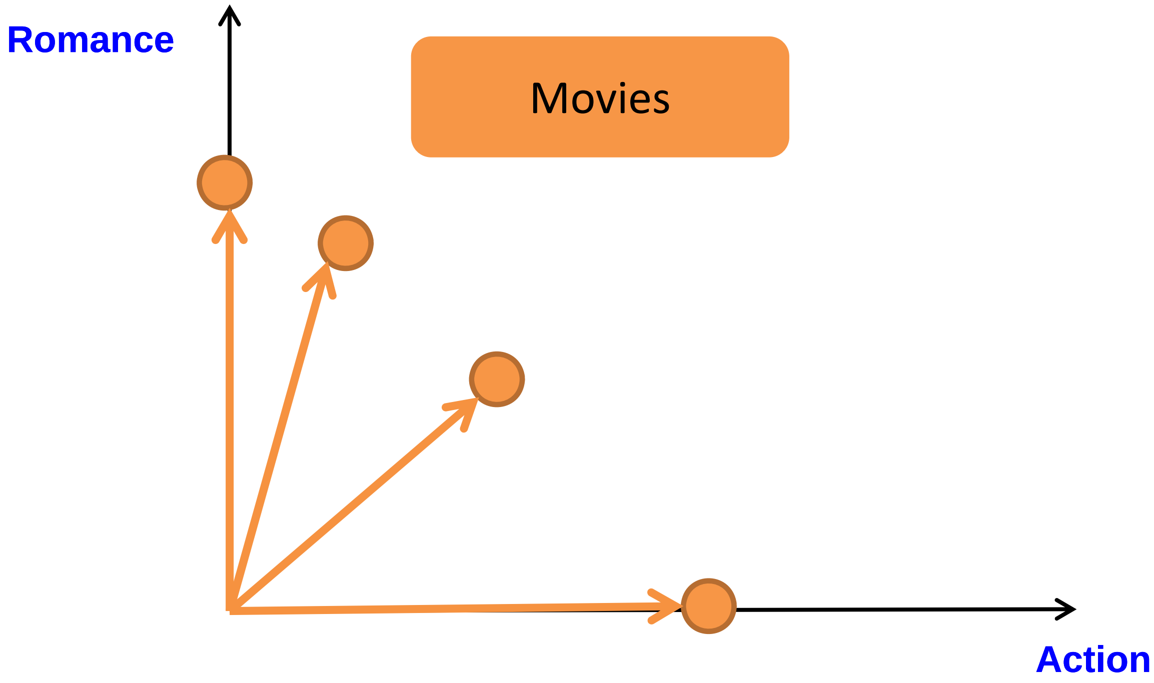


# Vector space model

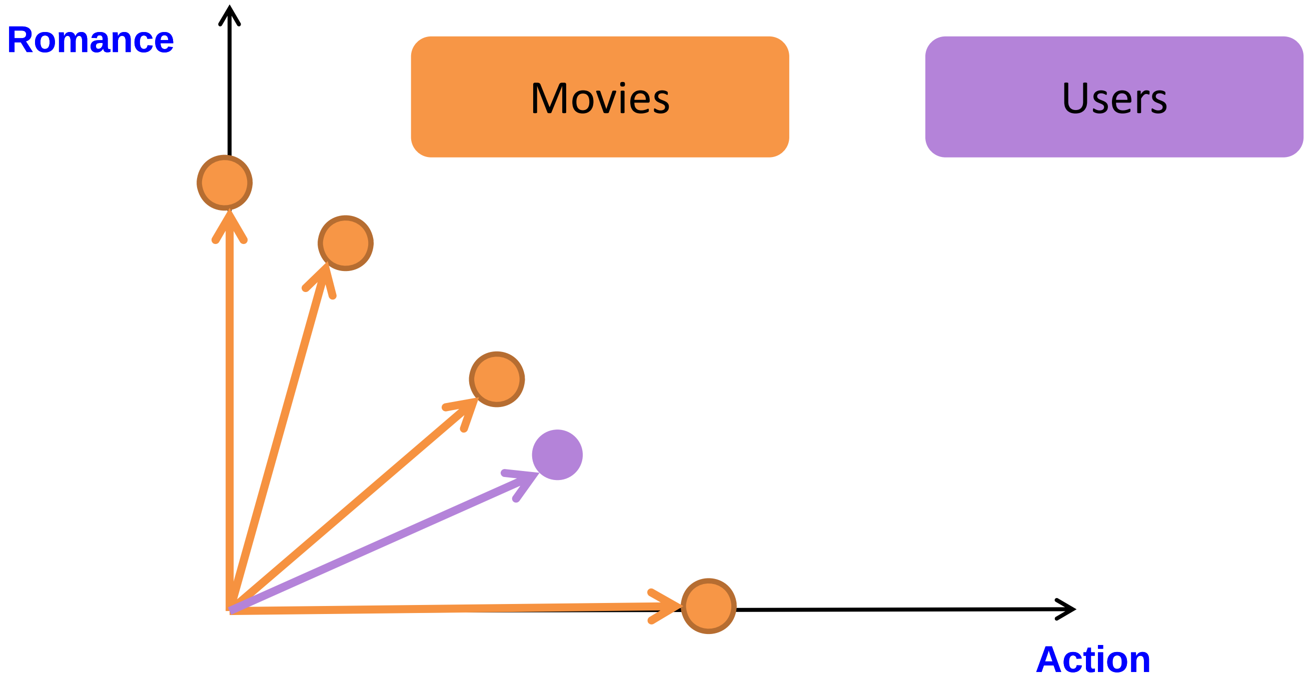
- It is an algebraic model for representing text documents (and any objects, in general) as vectors of identifiers.
- Originally created for queries, indexing
- Elements are represented as vectors

$$d_j = (w_{1,j}, w_{2,j}, \ldots, w_{t,j})$$

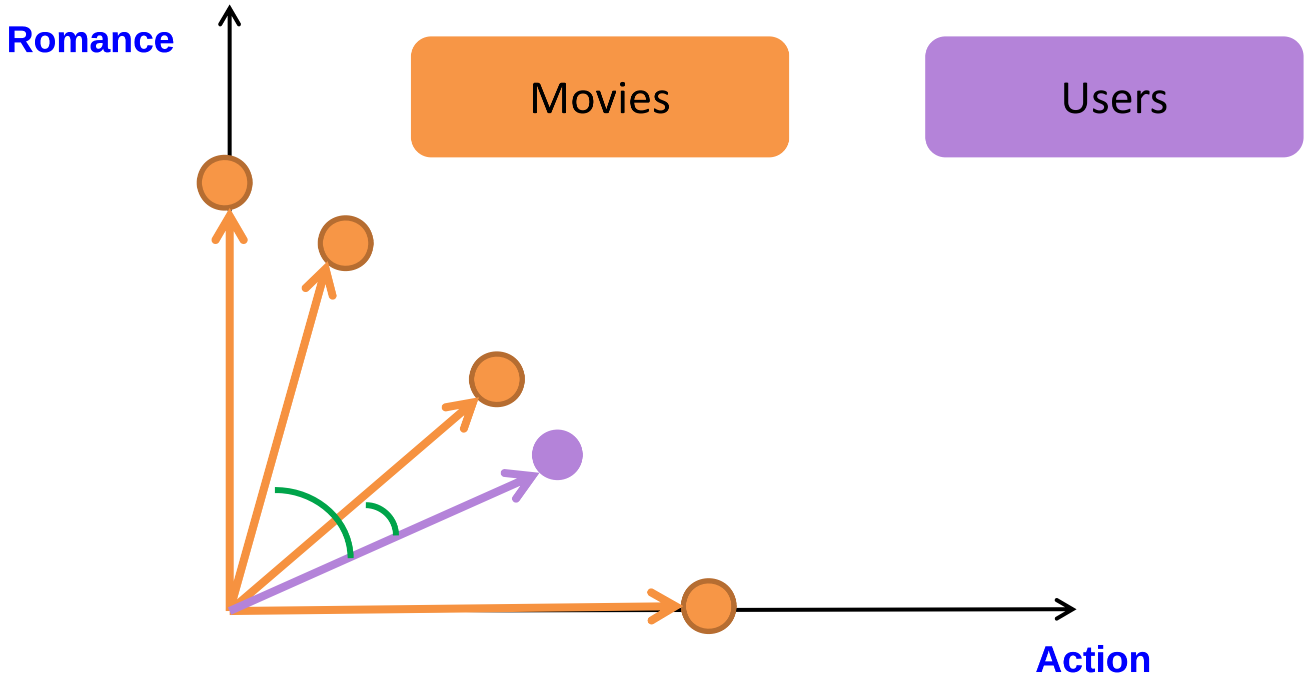
# Vector space model, items



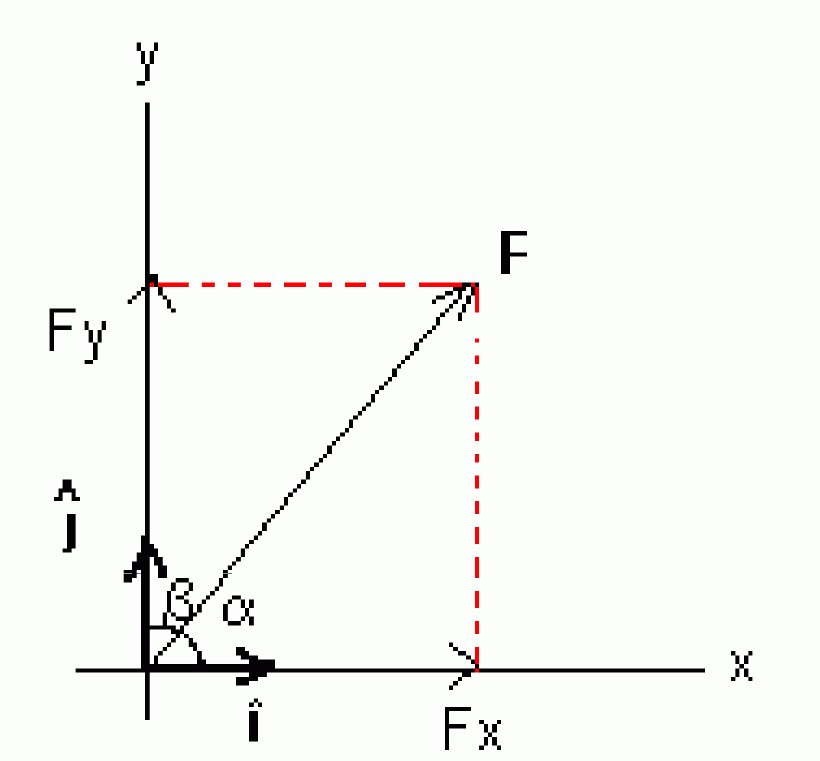
# Vector space model, users



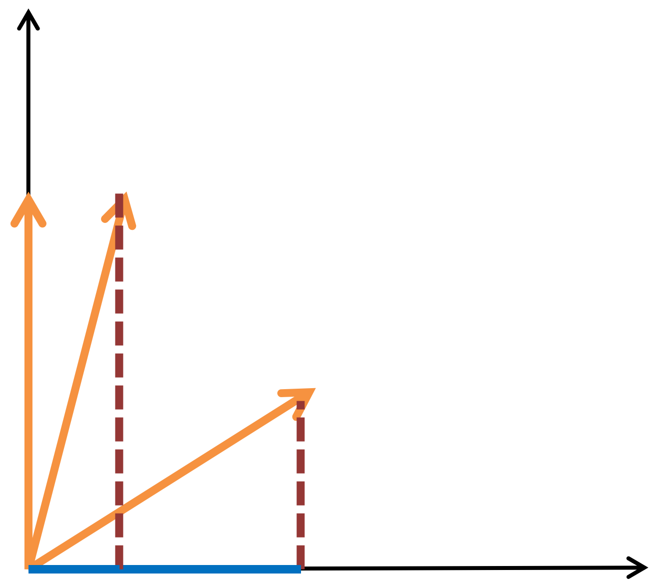
# Vector space model, Recommendation



# Vector space model, Recommendation




$$F_x = \| F \| \cos\alpha$$



**Remember**
Cos 0 = 1
Cos 90 = 0
Cos 180 = -1

# Vector space model, Recommendation

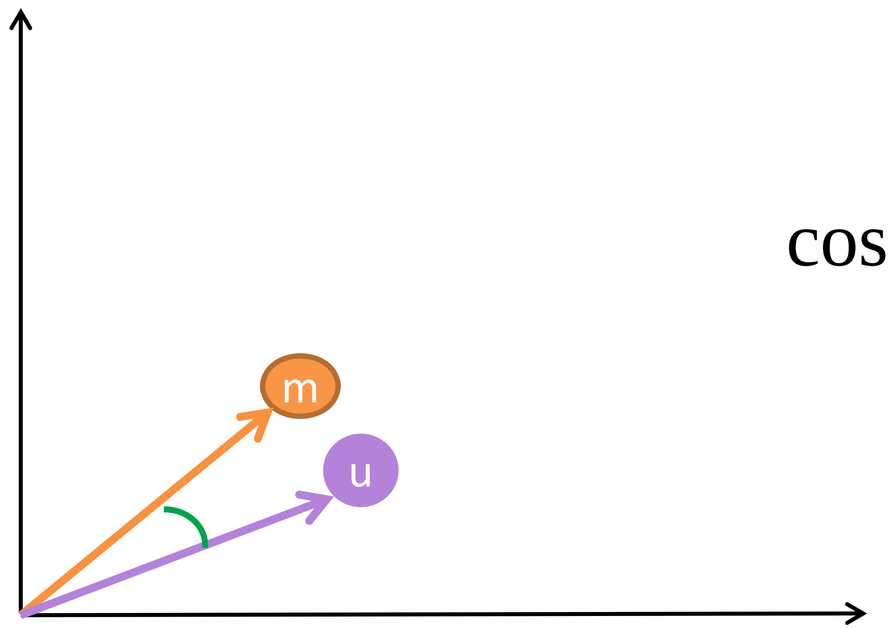
$$\cos\alpha = \frac{u \cdot v}{\|u\| \cdot \|v\|} = \frac{\sum_{i=1}^{n} u_i v_i}{\|u\| \cdot \|v\|}$$

Norm of the vector: Vector lenght

$$\|u\| = \sqrt{\sum_{j=1}^{n} u_j^2}$$

# Vector space model, Recommendation

- Prediction is the cosine of the angle between the two vectors (user profile, item)

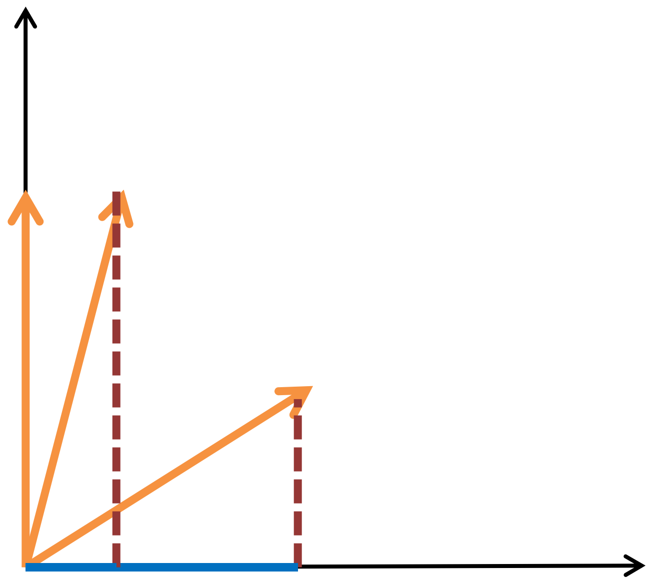



$$\cos\theta = \frac{m \cdot u}{\| m \| \cdot \| u \|}$$

# Vector space model, Recommendation

- Cosine
  - ranges between -1 and 1 (0 and 1 if all positive values in vectors)
  - closer to 1 is better.
- Adequate for Top-n

**Remember**
Cos 0 = 1
Cos 90 = 0
Cos 180 = -1



# Describing items

- Representing an item through a keyword vector:
  - Simple 0/1 (keyword applies or doesn't) Lacks intensity
  - Simple occurrence count Provides intensity
  - TFIDF, most commonly Provides intensity and distinctiveness
  - Other options
    - Document length

# Describing items

- Do we consider tags to be yes-or-no?
  - Actor (we don’t really get a measure for how much “Tom Hanks” a movie has)
  - Descriptive (is how often a tag is applied a proxy for how relevant/significant the feature is?)

Better count

- Do we care about IDF?
  - Actor (are infrequent actors more significant than stars?) Not adequate
  - Descriptive (is “prison scene” more significant than “car chase” or “romance”?)

Adequate

# Building user profiles

- A set of keywords that the user may like, dislike or not have an opinion on
- Infer from explicit and implicit user ratings are combined
  - Implicit. User actions: Read, buy, click
  - Explicit: Explicit user ratings
- Allowing the user to edit a profile can be also valuable

# Building user profiles, rankings

- Simply unary – aggregate profiles of items we rated without weights
- Unary with threshold – only put items above a certain rating into our profile (but all likes are equal)
- Weight, but positive only – higher weight for things with higher scores
- Weight, and include negative also – negative weight for low ratings

# Building user profiles, update profiles

- Don't recompute all each time
- Weight new/old similarly – keep track of total weight in profile and mix in new rating (linear combination)
  – Special case for changed rating; subtract old
- Mechanism for temporal decay.
  – Decay old profile and mix in new

# Building user profiles

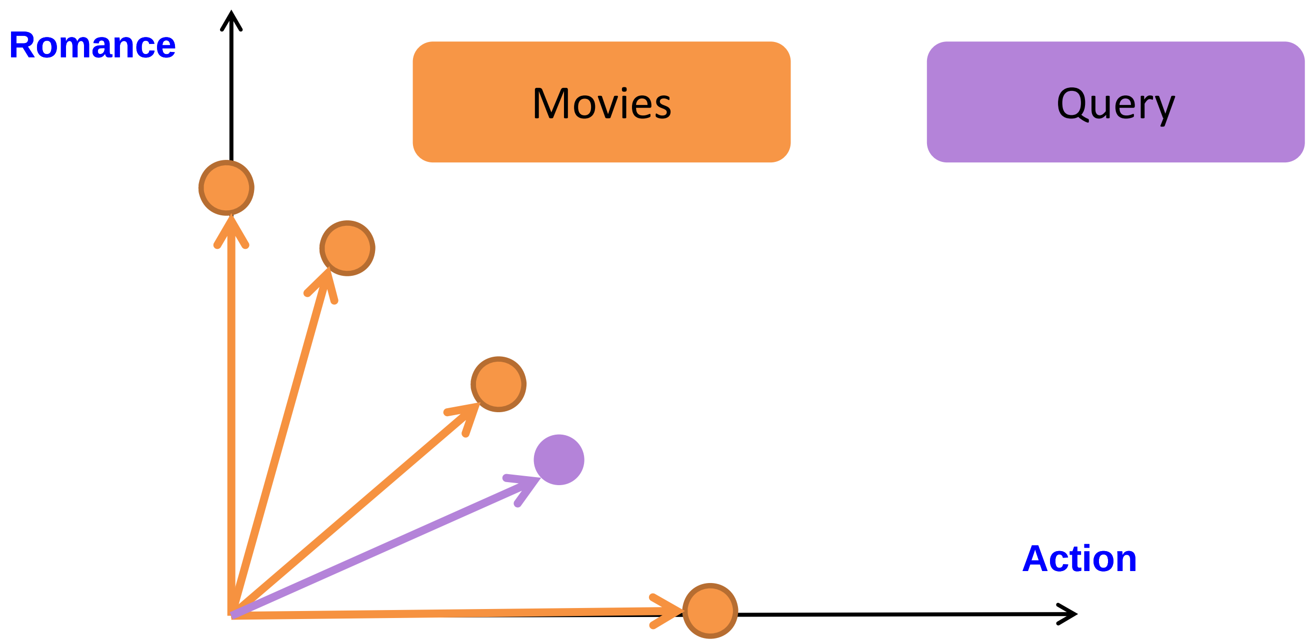
- Vector space model conflates liking with importance
  - Works well with query terms application
  - Not so well in others: I like ketchup a lot but I do not care much if it is in a dish I’m ordering

# Advantages/strengths

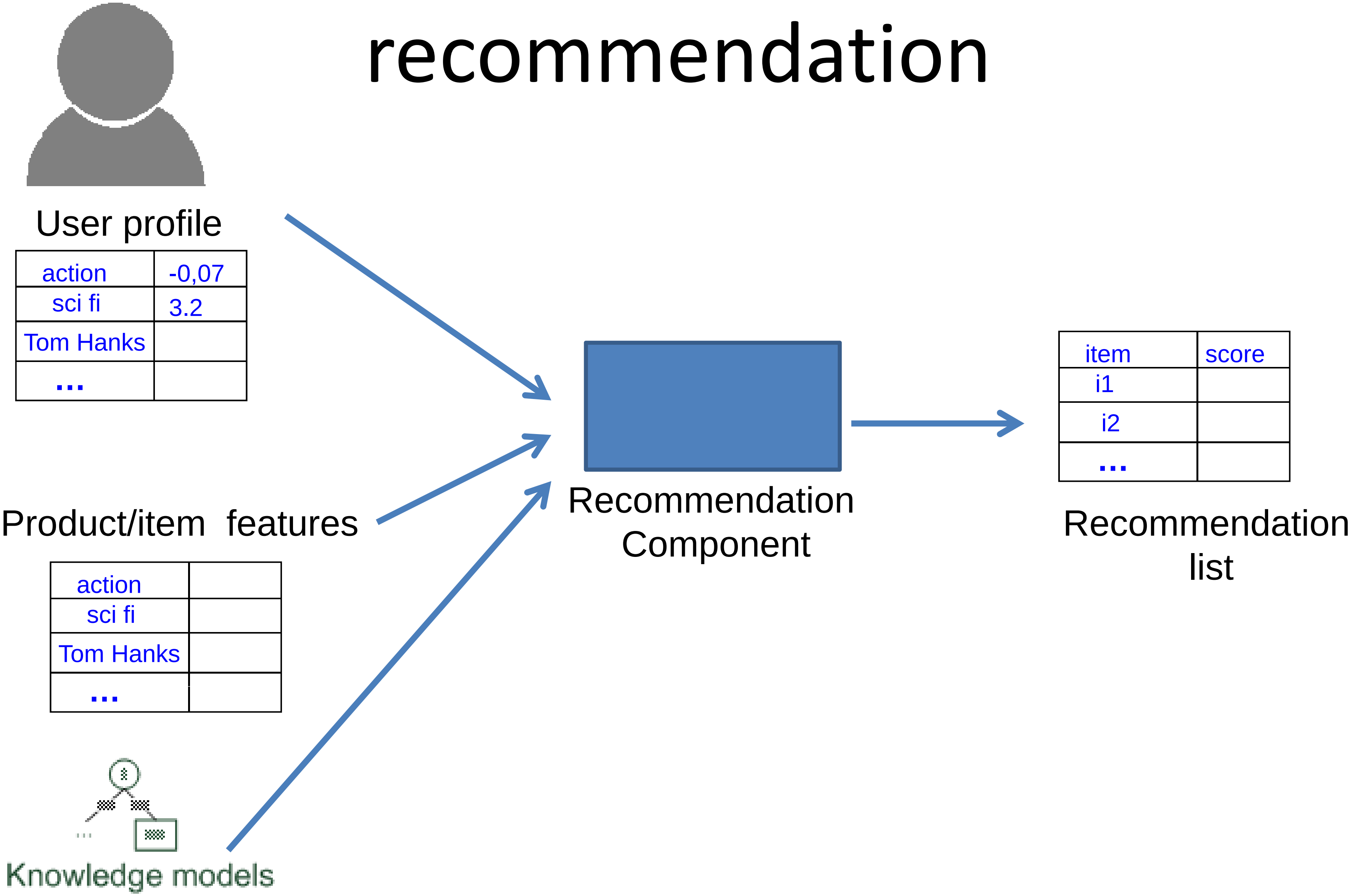
- Entirely content-based
- Understandable profile
- Easy computation
- Are capable of recommending items not yet rated by any user
- Flexibility to be integrated with query-based or case-based approaches

# Case-based recommendation

- Structure cases around a set of relevant attributes (e.g., camera price, zoom, pixels)
- Query based on an example or attribute query, and retrieve relevant cases



# Knowledge-based recommendation



# Challenges and limitations

- Figuring out the right weights and factors
- It cannot handle interdependencies
  - I like Sandra Bullock in Action movies, but Meg Ryan in Romantic Comedy movies
  - I like comedies with violence, and historical documentaries, but not historical comedies or violent documentaries

# References

- Vector space model
  - Salton, Wong, and Yang (1995) “A Vector Space Model for Automatic Indexing,” *CACM 18:11.*
  - http://en.wikipedia.org/wiki/Vector_space_model
- Ricci et.al. (2011).*Recommender Systems Handbook*. Chapter 3