

## Lecture 26

- Explaining Recommendations

---

IT492: Recommendation Systems (AY 2023/24) — Dr. Arpit Rana

## Recommendation Explanation

**An explanation of a recommendation is any content additional to the recommendation itself that justifies the recommended item to the user.**

*We recommend you the movie “A Beautiful Mind” because it has features ‘drama’ and ‘biography’ that you liked before.*

This textual description justifies the movie “A Beautiful Mind” to the user by means of its features ‘drama’ and ‘biography’ which she liked before.

# Goal of Explanation

Recommender systems provide explanations to -

- reveal how a recommender has reached its conclusions (**transparency**),
- help users to modify or correct the assumptions (**scrutability**),
- help users make better decisions (**effectiveness**),
- increase user trust in the system (**trust**),
- help users make decisions more quickly (**efficiency**),
- influencing user behaviour (**persuasiveness**), and
- improve user acceptance of recommendations (**satisfaction**).

## Goal of Explanation

**In addition to supporting end-users, explanations of recommendations may have a role in issues such as:**

- detecting shilling attacks;
- detecting bias and discrimination; and
- contesting algorithmic decisions on personal data  
(as allowed for in government legislation such as GDPR in Europe (2016) and PDPB in India (2019))

# Characterizing Explanations of Recommendations

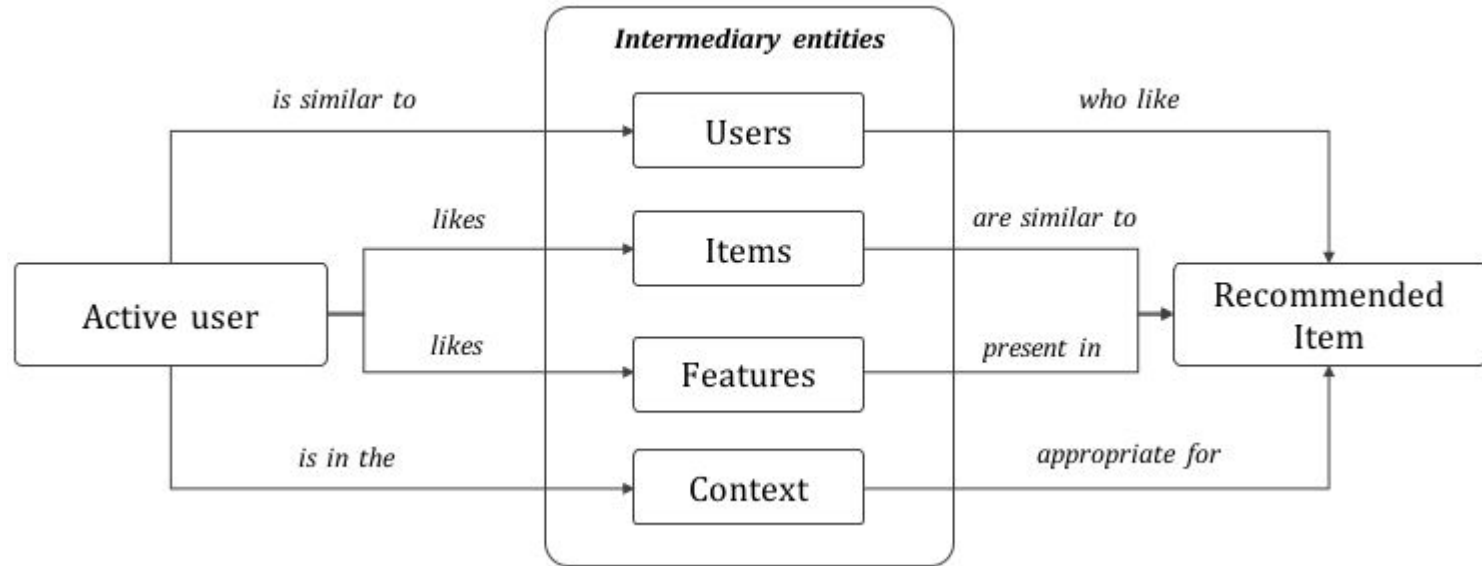
**Explanations of recommendations can be characterized in a variety of different ways such as:**

- the type of knowledge they use (e.g. user demographics, item descriptions, etc.)
- their fidelity to the recommender (i.e. white-box vs. black-box), and
- their role in producing recommendations.

## Based on their Type of Knowledge

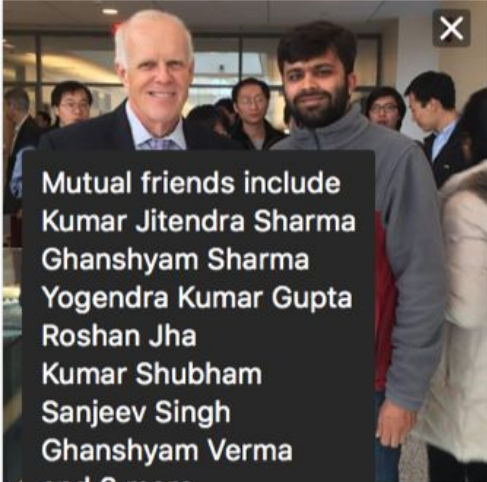
Explanations of recommendations often relate the recommended item to the user through intermediary entities,

e.g. other users, other items, item features, or context




# Based on their Type of Knowledge


## User-based Explanations



Mutual friends include  
Kumar Jitendra Sharma  
Ghanshyam Sharma  
Yogendra Kumar Gupta  
Roshan Jha  
Kumar Shubham  
Sanjeev Singh  
Ghanshyam Verma  
and 6 more...


7 mutual friends

 Add Friend



**Data Science &  
Machine Learning**

26,852 Members

 7 connections have  
joined

[Join](#)



**Science Foundation  
Ireland**

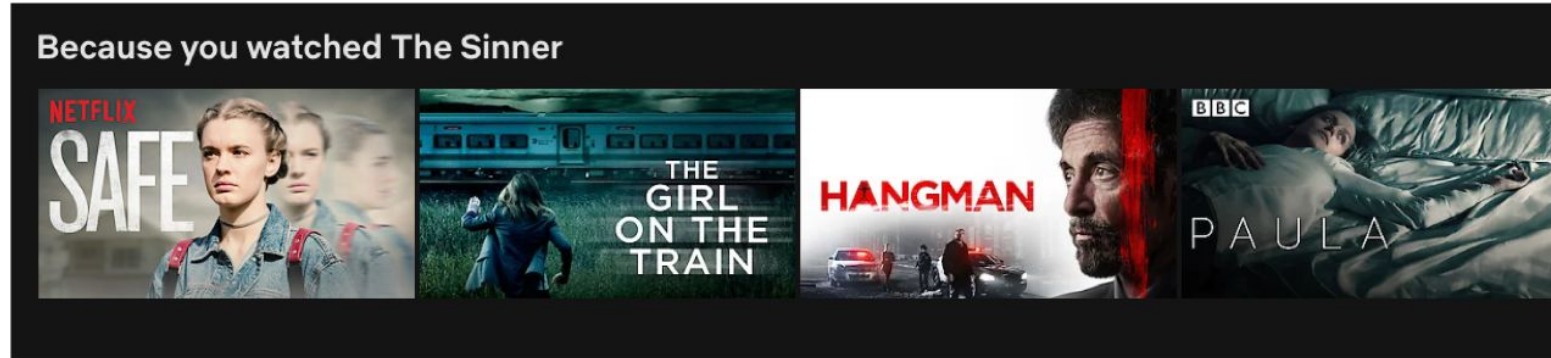
7,837 Followers

People in your community  
follow

[Follow](#)

# Based on their Type of Knowledge

## Item-based Explanations



### Because you viewed

Research Scientist at Amazon

[See all](#)



**Data Scientist**  
LinkedIn  
Bangalore



**Research Fellow**  
Microsoft  
Bangalore, IN



**Deep Learning Data Scientist**  
Intel Corporation  
Bangalore, IN



**Research Analyst**  
WAVTEQ  
Bengaluru, Karnataka, India



## Based on their Type of Knowledge

### Feature-based Explanations:

- *Attribute-value pairs*  
A movie based on the user's most preferred actor, genres and director
- *Item content*  
News, books, articles or blogs based on the keywords extracted from their textual content
- *User-generated tags*  
Items using tags that users assign to the items
- *User-reviews*  
Items using information extracted from user reviews
- *Linked data*  
Items using linked open data on DBpedia

## Based on their Type of Knowledge

### **Contextual Explanations:**

- Time, location, weather, or companions, can influence how a person perceives a product or service.
- Such observable contextual factors are used to explain the recommended items to the user.

## Based on their Fidelity

In Artificial Intelligence in general, explanations are sometimes categorized as

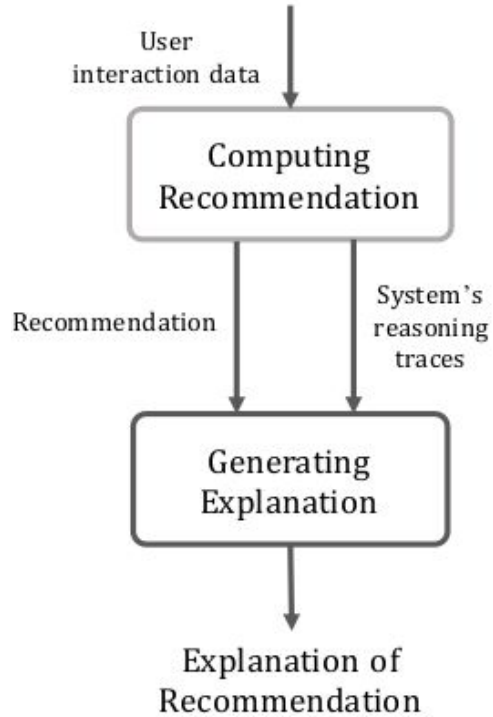
- **White-box (a.k.a. *model-based*)**

These are built from traces of the system's reasoning, e.g. content-based methods

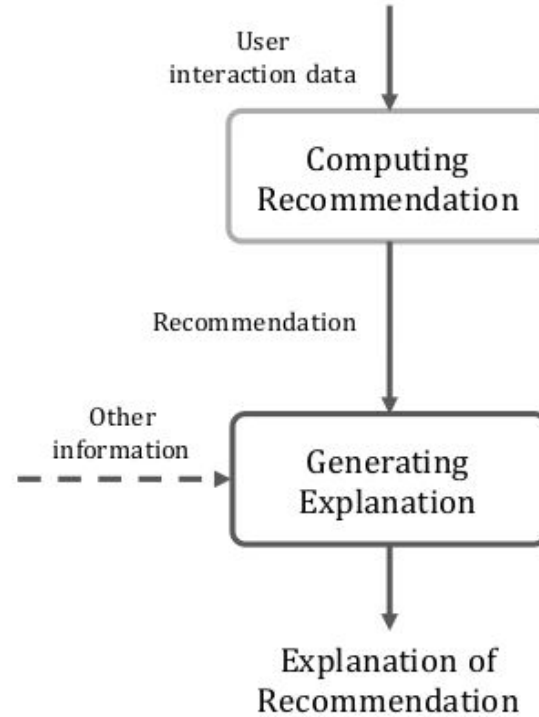
- **Black-box (a.k.a. *model-agnostic*)**

These explanations make no use of knowledge of how the system produced its decision, e.g. matrix factorization for recommendation

## Based on their Fidelity

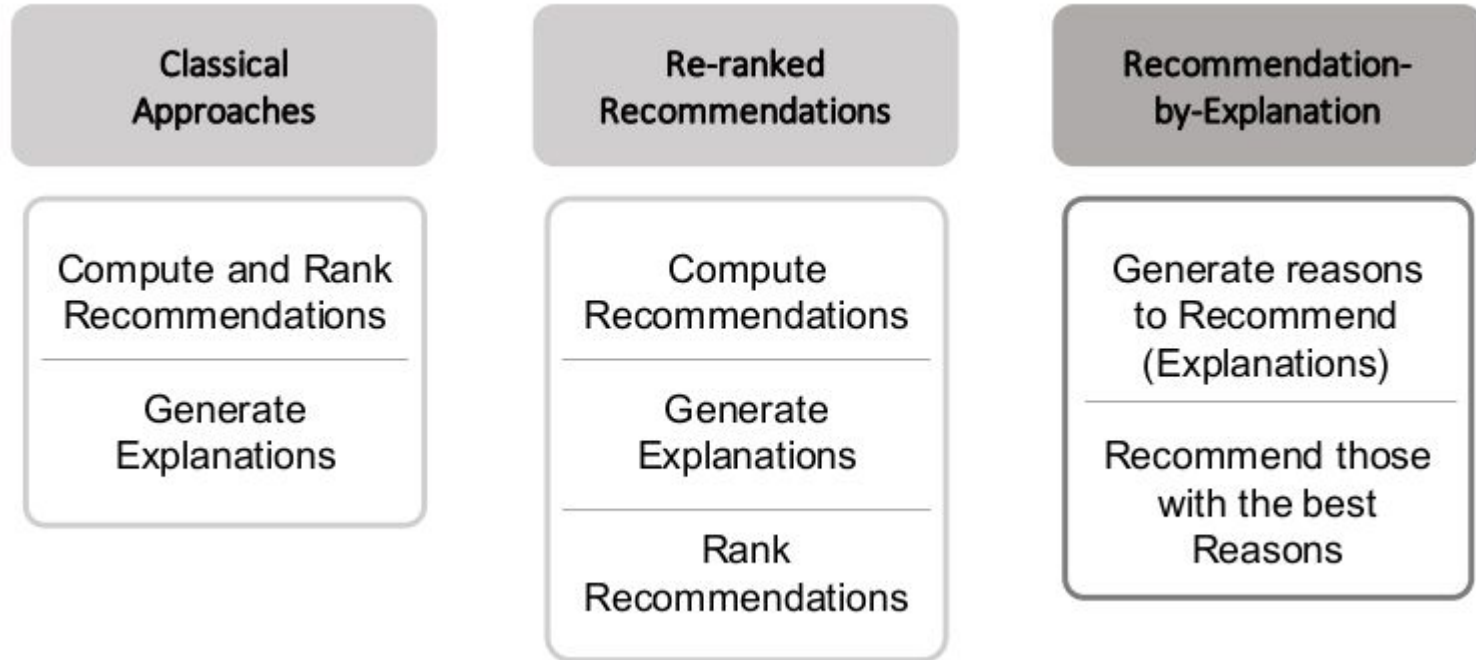


**(a) White-box explanations**



**(b) Black-box explanations**

## Based on their Role in Producing Recommendations



# Evaluating Explanations of Recommendations

## Explanations are user-centric.

- Offline experiments are of very limited use; for example, we can measure the size of a system's explanation (e.g. how many items or features they contain).
- To evaluate the subjective perception of the users and their impact on user behaviour really requires either user trials or online evaluation with a deployed system.

# Evaluating Explanations of Recommendations

## Satisfaction vs. Promotion [Bilgic & Mooney, 2005]

- A user is initially asked to rate a recommendation where she is given only the explanation and not the identity of the item. This is called the explanation-rating.
- The user is asked later to re-rate the recommended item in the case where she is not given the explanation but she is given information about the item, including its identity. This is called the actual-rating.
- Explanations are effective if both the ratings are close to each other.
- Explanations are persuasive if explanation-rating is higher than the actual-rating.

# Explanations Dimensions

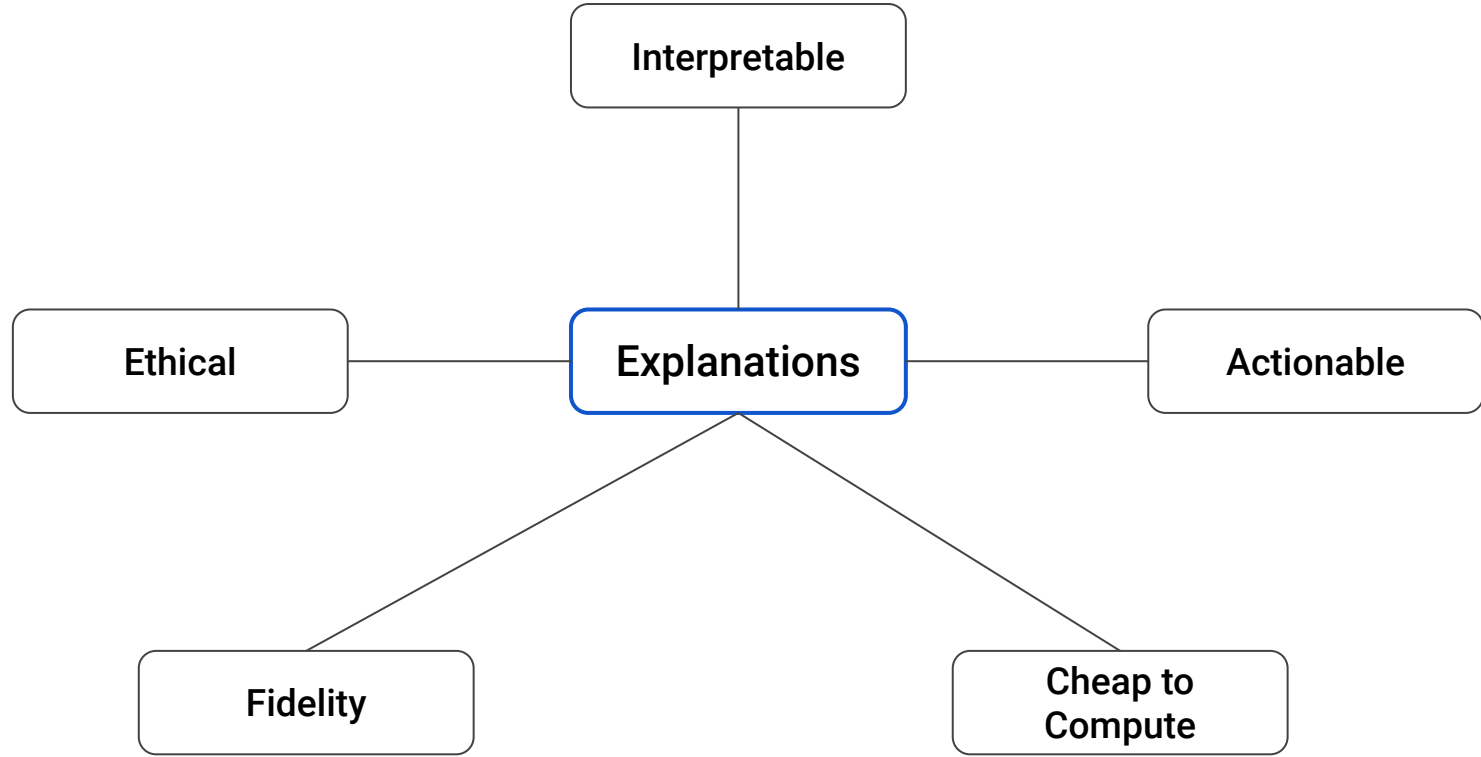


Diagram is taken from Derek Bridge's talk on Fidelity vs. Interpretability.



## Next Lecture

- Sequential Recommendations