



Carnegie Mellon University
Tepper School of Business

46-886: Machine Learning Fundamentals

Amr Farahat

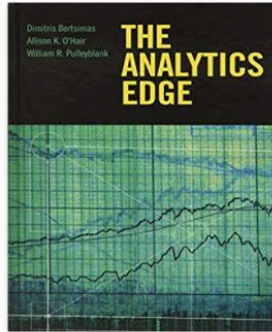
Recommender Systems: Introduction

Much of this slide deck is derived/borrowed from course material
I've co-taught at MIT



This work is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Amazon Recommendations



See all 2 images

The Analytics Edge Hardcover – 2016

by [Dimitris Bertsimas](#) (Author), [Allison O'Hair](#) (Author), [William Pulleyblank](#) (Author)

★★★★☆ 5 customer reviews

[See all 4 formats and editions](#)

Hardcover
\$110.00

6 Used from

\$209.18

3 New from

\$110.00

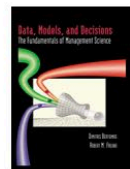
The Analytics Edge provides a unified, insightful, modern and entertaining treatment of analytics. The book covers the science of using data to build models, improve decisions, and ultimately add value to institutions and individuals.

The philosophical underpinnings of the book are that real world problems are
[Read more](#)

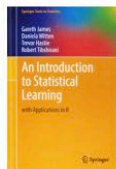
[Report incorrect product information.](#)

Discover the best books
of 2019 so far

Customers who viewed this item also viewed



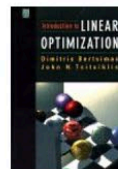
Data, Models, and Decisions: The Fundamentals of...
by [Dimitris Bertsimas](#)
★★★★☆ 14
Hardcover
61 offers from
\$36.64



An Introduction to Statistical Learning: with Applications in R...
by [Gareth James](#)
★★★★☆ 214
#1 Best Seller in
Mathematical Physics
Hardcover
\$48.61 ✓prime



Data Science for Business: What You Need to Know about Data Mining and...
by [Foster Provost](#)
★★★★☆ 198
Paperback
\$32.80 ✓prime



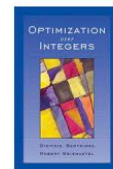
Introduction to Linear Optimization (Athena Scientific Series in...)
by [Dimitris Bertsimas](#)
★★★★☆ 39
Hardcover
\$89.00 ✓prime



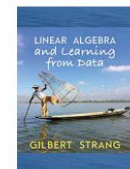
Intro Stats (4th Edition)
by [Richard D. De Veaux](#)
★★★★☆ 103
Hardcover
\$170.99 ✓prime



Discovering Statistics Using R
by [Andy Field](#)
★★★★☆ 95
Paperback
\$70.46 ✓prime



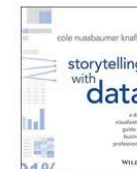
Optimization Over Integers
by [Dimitris Bertsimas](#)
★★★★☆ 2
Hardcover
\$110.00 ✓prime



Linear Algebra and Learning from Data
by [Gilbert Strang](#)
★★★★☆ 6
Hardcover
\$67.94



Applied Predictive Modeling
by [Max Kuhn](#)
★★★★☆ 83
Hardcover
\$65.97 ✓prime



Storytelling with Data: A Data Visualization Guide for Business Professionals
by [Cole Nussbaumer...](#)
★★★★☆ 335
#1 Best Seller in
Business Mathematics
Paperback
\$27.55 ✓prime

Types of Recommendation Systems

Demographic filtering

- Based on user demographics

The Facebook logo, consisting of the word "facebook" in a blue, lowercase, sans-serif font.

Content-based filtering

- Based on item features and characteristics

The Pandora logo, featuring the word "PANDORA" in a blue, serif font, with "internet radio" in a smaller, lowercase, sans-serif font below it.

Collaborative filtering

- Based on user-item interaction data from all users and items

The Netflix logo, featuring the word "NETFLIX" in a white, bold, sans-serif font on a red rectangular background.

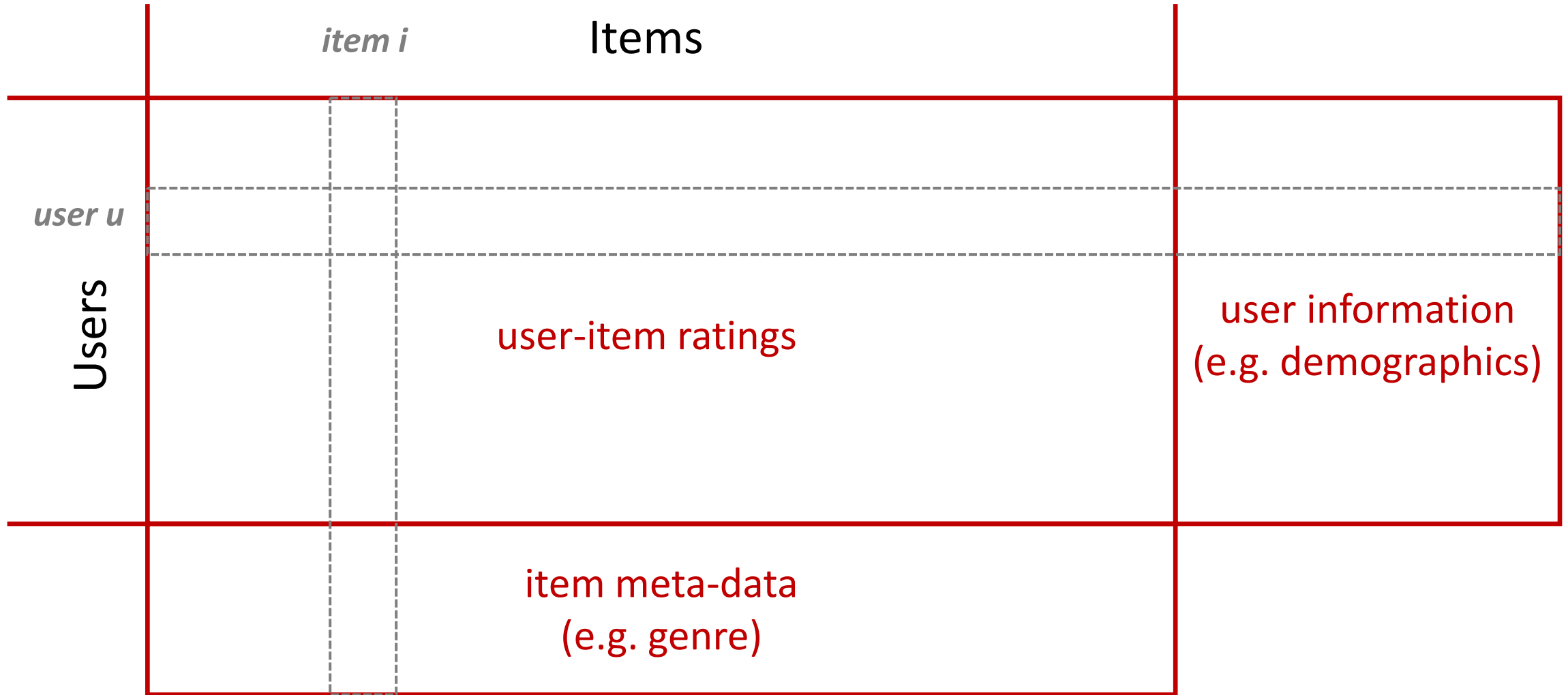
Matching

- Based on compatibility: recommend matches that will “work” both ways

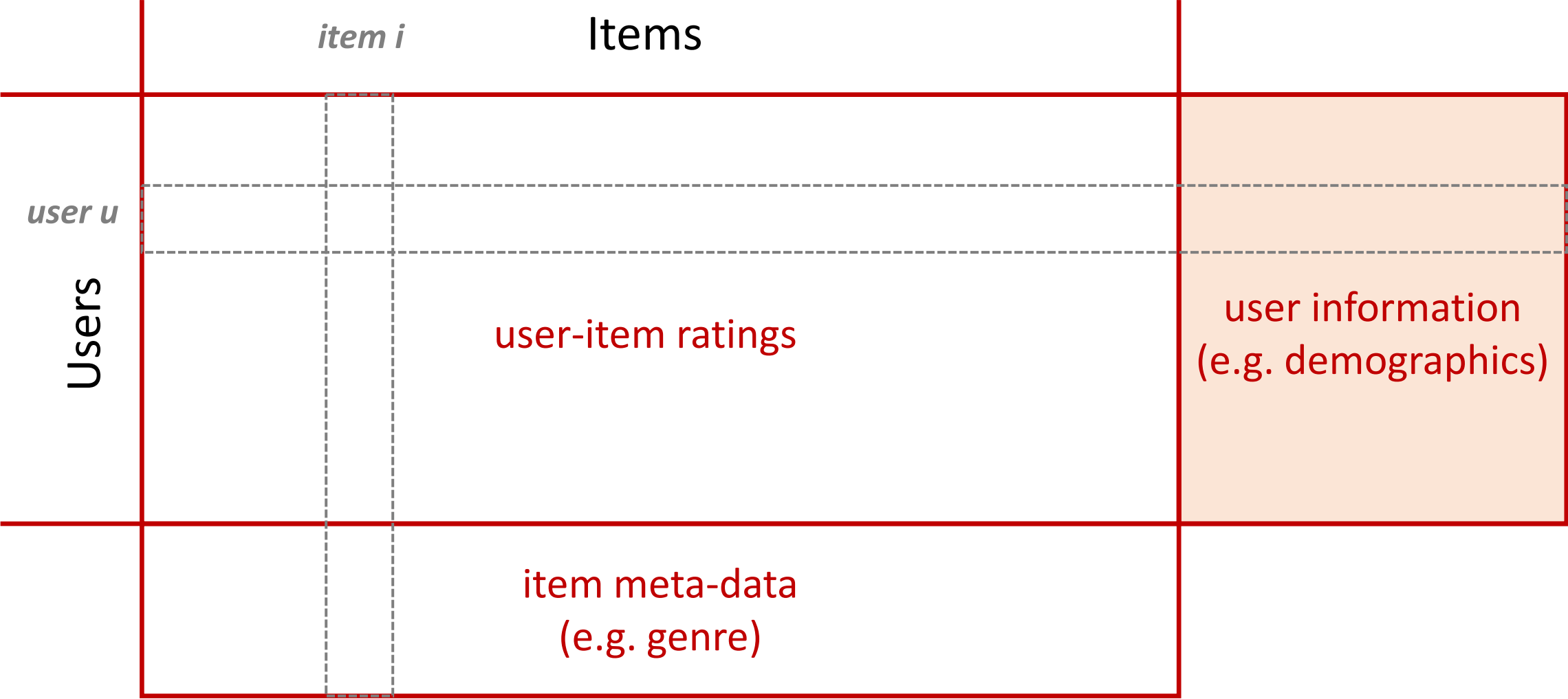
The eharmony logo, featuring a stylized heart icon made of four colored segments (blue, green, red, yellow) above the word "eharmony" in a teal, lowercase, sans-serif font.

But there are lots of hybrid and bespoke approaches out there.

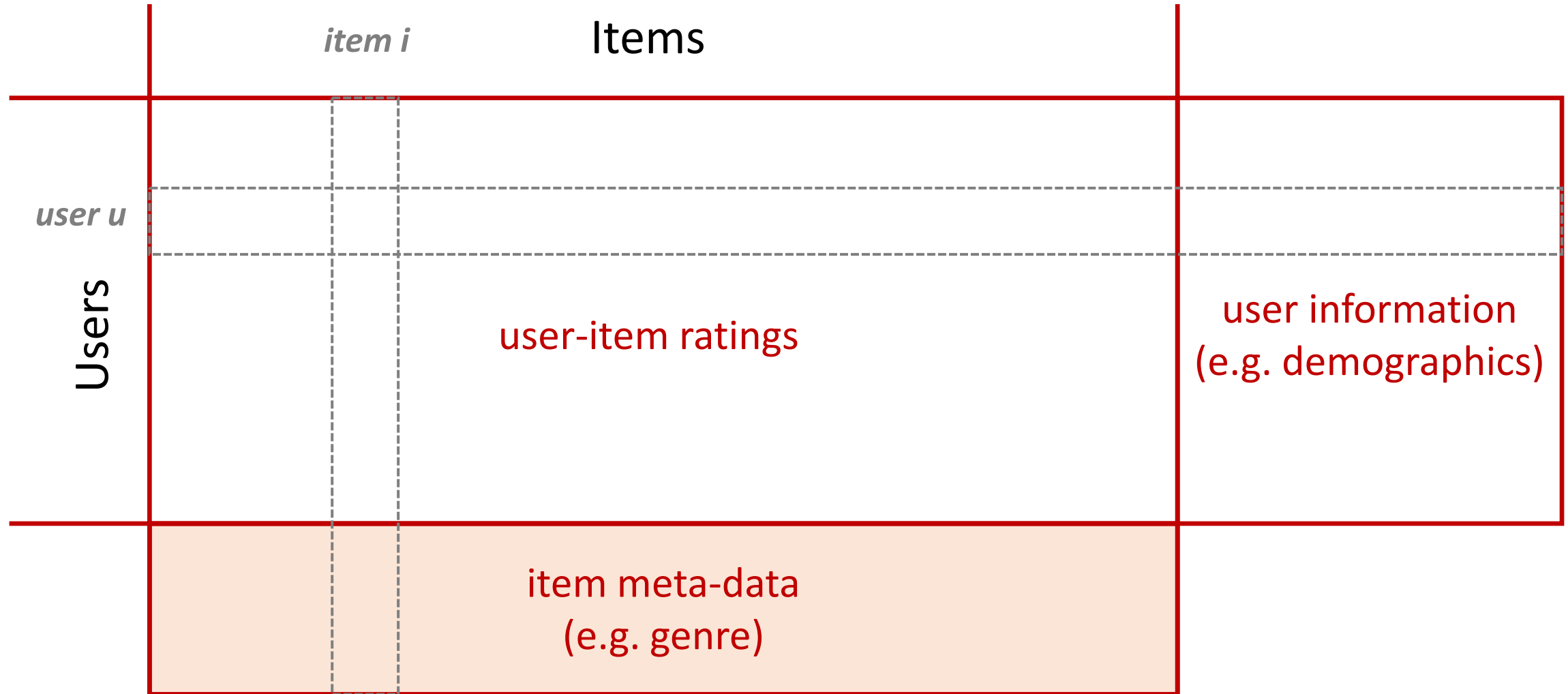
A Broad View of Data Underlying Recommender Systems



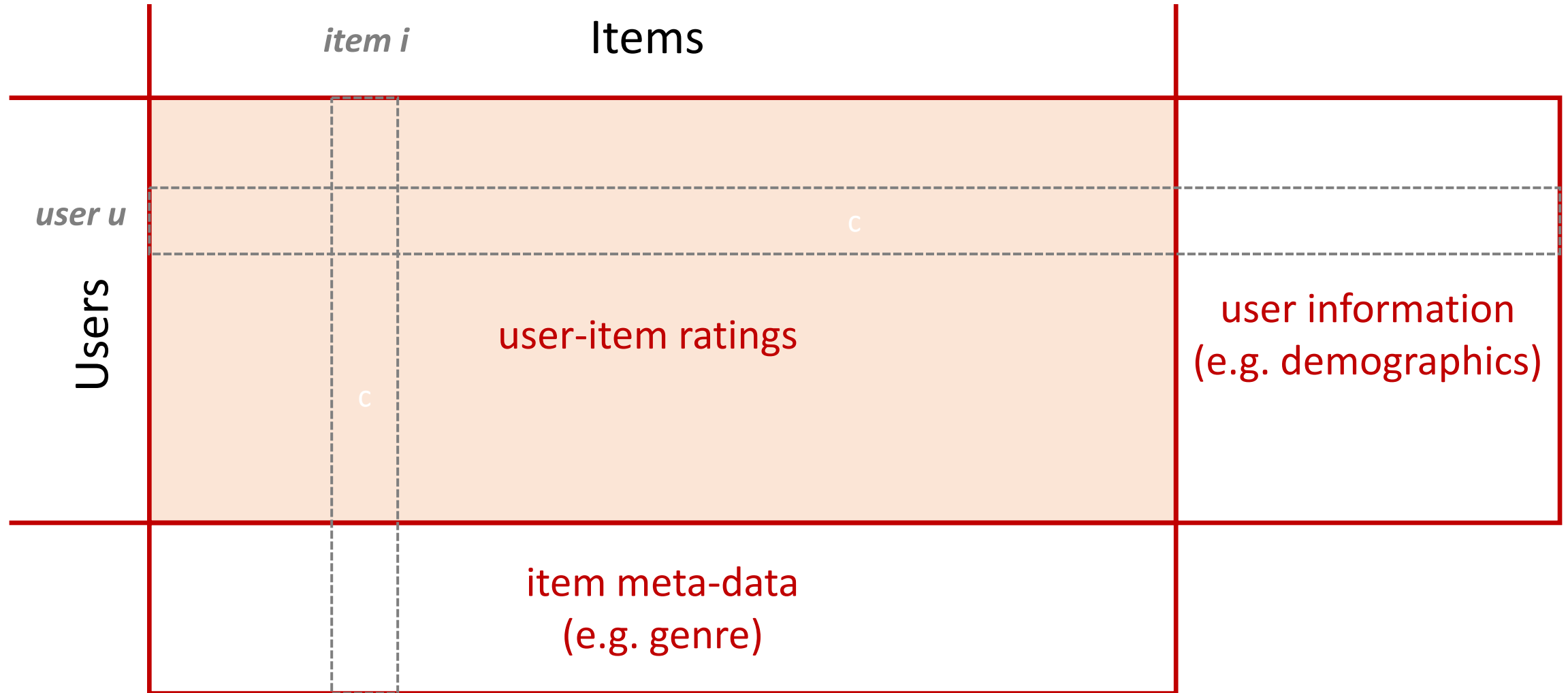
Demographic Filtering



Content-Based Filtering



Collaborative Filtering



An Important Practical Consideration – The Cold Start Problem

- For a new user, what do we recommend?
 - If you have user metadata (e.g., demographics, device, location) data, you can make recommendations that are popular for the average user in that ‘metadata bucket’
 - You can recommend items popular across the entire user base and rapidly adapt based on how the user reacts to those recommendations
- For a new item, what do we do?
 - If you have item metadata (e.g, genre, director, lead actor for movies), you can assign the average rating of the ‘metadata bucket’
 - Identify ‘sibling items’ and ‘borrow’ the sibling’s ratings as a starting point