CHAPTER 7
WEB BASED INFORMATION SYSTEM AND NAVIGATION

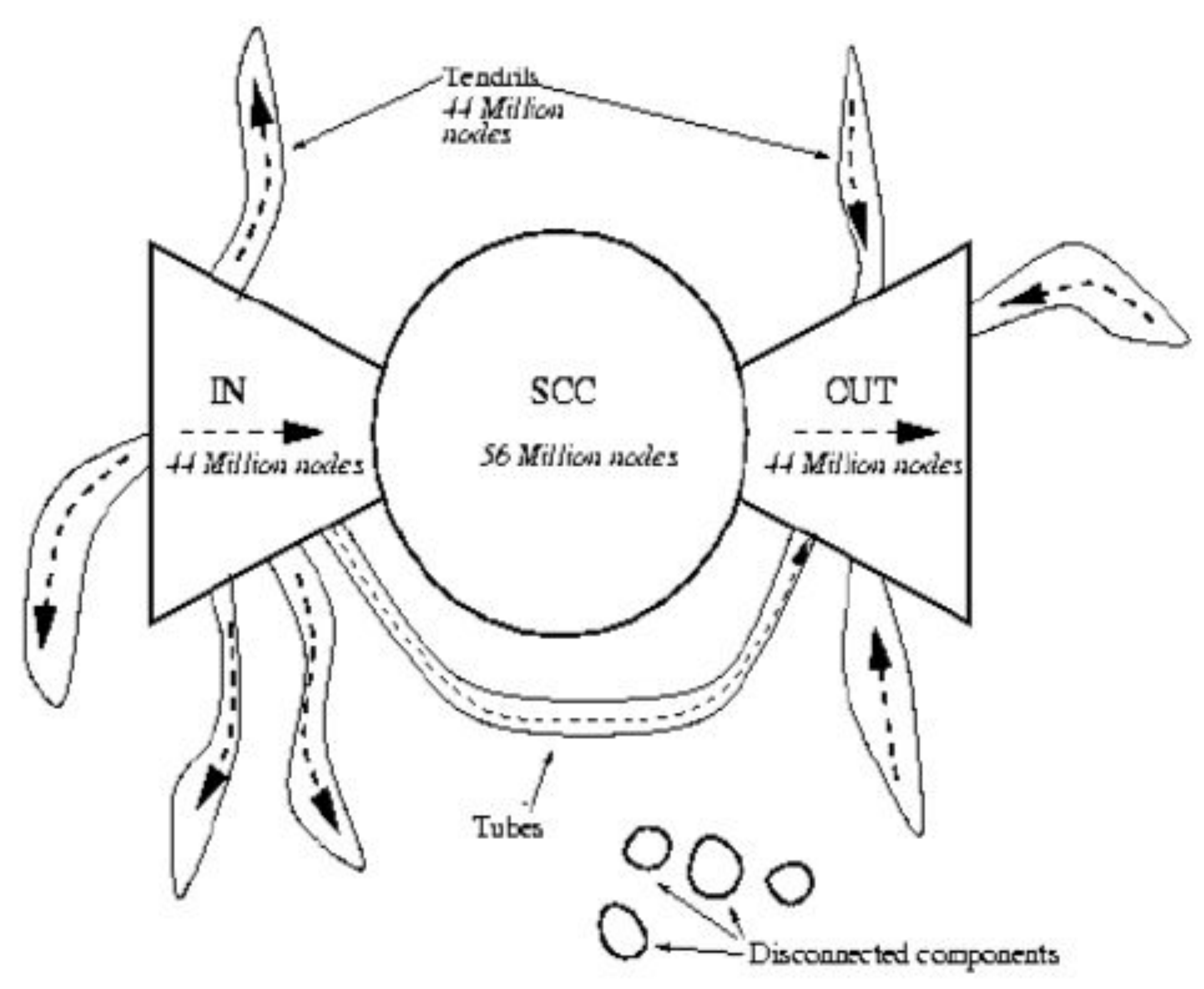
Syllabus

- The structure of the web
- Link Analysis
- Searching the web
- Navigating the web
- Web uses mining
- Collaborative filtering
- Recommender systems
- Collective intelligence

7.1Structure of the Web

- The Web is the common name for the World Wide Web, a subset of the Internet consisting of the pages that can be accessed by a Web browser.
- Many people assume that the Web is the same as the Internet, and use these terms interchangeably.
- It has no engineered architecture or structure, but rather a virtual network of content and hyperlinks and consist a large, strongly connected core in which every page can reach every other by a path of hyperlinks.
- The structure of Web is explained below:

DIFFERENT NODES ARE UPSTREAM, DOWNSTREAM AND TENDRIL NODES.



Different nodes:

- □In the SCC
- □In the “inbound” part
- □In the “outbound” part
- □Tendrils
- □Disconnected nodes

- In 1999, after the Web had been growing for the better part of a decade, Andrei Broder and his colleagues set out to build a global map of the Web, using strongly connected components (SCC).
- The Bow-Tie Structure-The step include to position all the remaining SCCs in relation to the giant one.
- This involves classifying nodes by their ability to reach and be reached from the giant SCC. The first two sets in this classification are the following.
- **1.IN**: a
- Nodes that can reach the giant SCC but cannot be reached from it – i.e., nodes that are “upstream” of it.
- **2.OUT**:
- Nodes that can be reached from the giant SCC but cannot reach it –i.e., nodes are “downstream” of it.

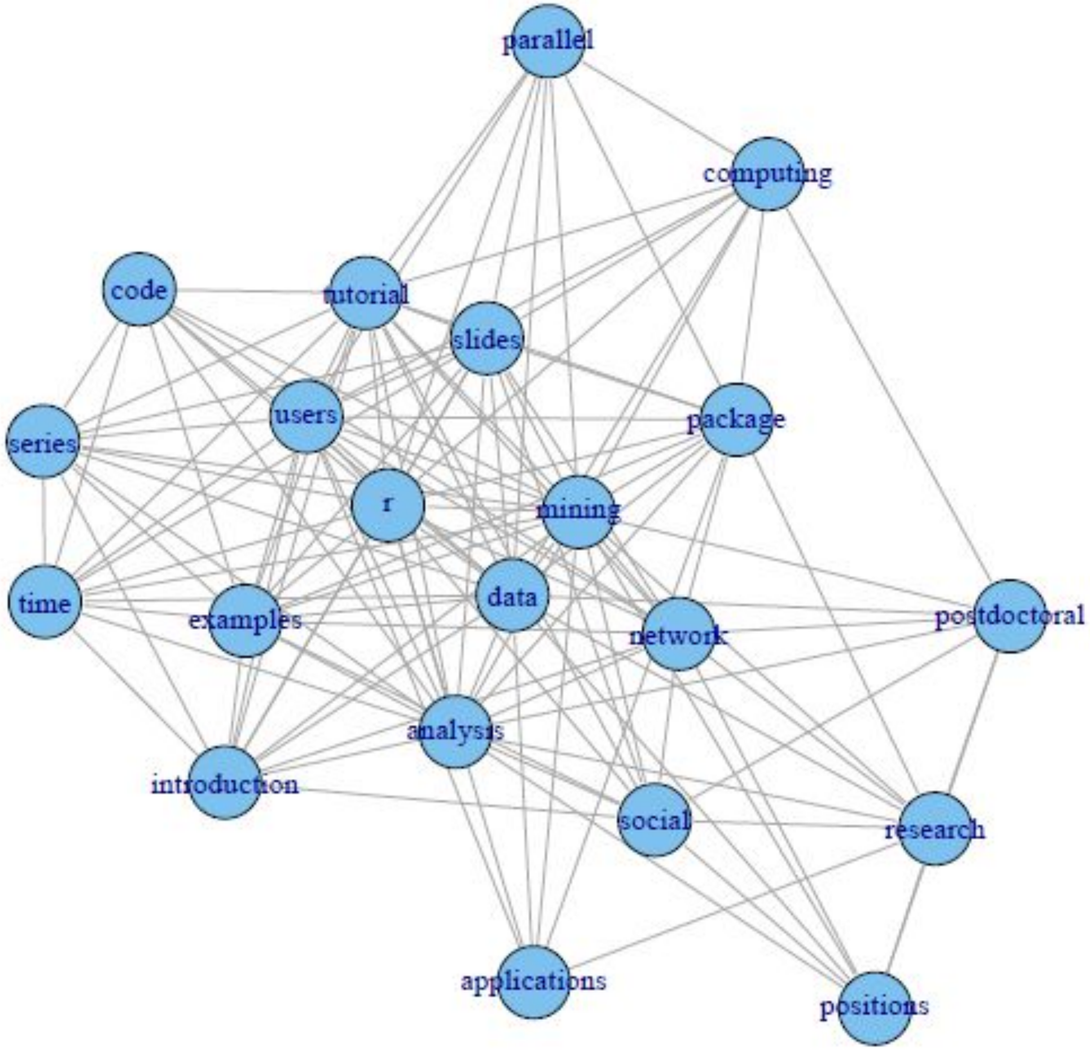
- **3.**The strongly connected components or disconnected components of an arbitrary directed graph form a partition into sub-graphs that are themselves strongly connected.
- **4.Tendrils**:
- The “tendrils” of the bow-tie consist of (a) the nodes reachable from IN that cannot reach the giant SCC, and (b) the nodes that can reach OUT but cannot be reached from the giant SCC.

- **5.Disconnected:**
- Finally, there are nodes that would not have a path to the giant SCC even if we completely ignored the directions of the edges. These belong to none of the preceding categories.
- A central core contains pages between which users can surf easily. Another large cluster, labelled ‘in’, contains pages that link to the core but cannot be reached from it.
- These are often new pages that have not yet been linked to. A separate ‘out’ cluster consists of pages that can be reached from the core but do not link to it, such as corporate websites containing only internal links.
- Other groups of pages, called ‘tendrils’ and ‘tubes’, connect to either the in or out clusters, or both, but not to the core, whereas some pages are completely unconnected.
- To illustrate this structure, the researchers picture the web as a plot shaped like a bow tie with finger-like projections.

7.2Link Analysis

- Link analysis is a data analysis technique used in network theory that is used to evaluate the relationships or connections between network nodes.
- These relationships can be between various types of objects (nodes), including people, organizations and even transactions.
- Link analysis is essentially a kind of knowledge discovery that can be used to visualize data to allow for better analysis, especially in the context of links, whether Web links or relationship links between people or between different entities.
- Link analysis is often used in search engine optimization as well as in intelligence, in security analysis and in market and medical research.

- A means of finding authoritative, relevant sources on the Web has proven useful in the design of improved search engines
- A data analysis technique used to evaluate relationships (connections) between nodes
- leading to improved methods for accessing and understanding the available information
- Web pages can be defined as hubs and authorities.
- Characteristic patterns for to identify communities of pages on the same topic


parallel
computing
code
tutorial
slides
users
series
package
r
mining
time
data
examples
network
postdoctoral
analysis
introduction
social
research
applications
positions

7.3Searching the Web

- A web search engine is a software system designed to search for information on the World Wide Web. Searching in the web is done by search engine
- A search engine operates in the following order:
- ■Web crawling
- ■Indexing
- ■Searching

- Web search engines work by storing information about many web pages, which they retrieve from the HTML markup of the pages. These pages are retrieved by a Web crawler (sometimes also known as a spider) — an automated Web crawler which follows every link on the site. The site owner can exclude specific pages by using robots.txt.
- □The search engine then analyzes the contents of each page to determine how it should be indexed (for example, words can be extracted from the titles, page content, headings, or special fields called meta tags).
- □Data about web pages are stored in an index database for use in later queries. A query from a user can be a single word.
- □The index helps find information relating to the query as quickly as possible.
- □Some search engines, such as Google, store all or part of the source page (referred to as a cache) as well as information about the web pages, whereas others, such as AltaVista, store every word of every page they find.

- □This cached page always holds the actual search text since it is the one that was actually indexed, so it can be very useful when the content of the current page has been updated and the search terms are no longer in it.
- □This problem might be considered a mild form of linkrot, and Google's handling of it increases usability by satisfying user expectations that the search terms will be on the returned webpage.
- □Increased search relevance makes these cached pages very useful as they may contain data that may no longer be available elsewhere.
- □Most Web search engines are commercial ventures supported by advertising revenue and thus some of them allow advertisers to have their listings ranked higher in search results for a fee. Search engines that do not accept money for their search results make money by running search related ads alongside the regular search engine results. The search engines make money every time someone clicks on one of these ads.

7.4Web Navigation

- Web navigation refers to the process of navigating a network of information resources in the World Wide Web, which is organized as a hypertext or hypermedia.
- □The user interface that is used to do so is called a web browser.
- □A central theme in web design is the development of a web navigation interface that maximizes usability.
- □A website's overall navigational scheme includes several navigational pieces such as global, local, supplemental, and contextual navigation; all of these are vital aspects of the broad topic of web navigation.

- Hierarchal navigation systems are vital as well since it is the primary navigation system.
- It allows for the user to navigate within the site using levels alone, which is often seen as restricting and requires additional navigation systems to better structure the website.
- The global navigation of a website, as another segment of web navigation, serves as the outline and template in order to achieve an easy maneuver for the users accessing the site, while local navigation is often used to help the users within a specific section of the site.
- All these navigational pieces fall under the categories of various types of web navigation, allowing for further development and for more efficient experiences upon visiting a webpage.

7.5Web Mining

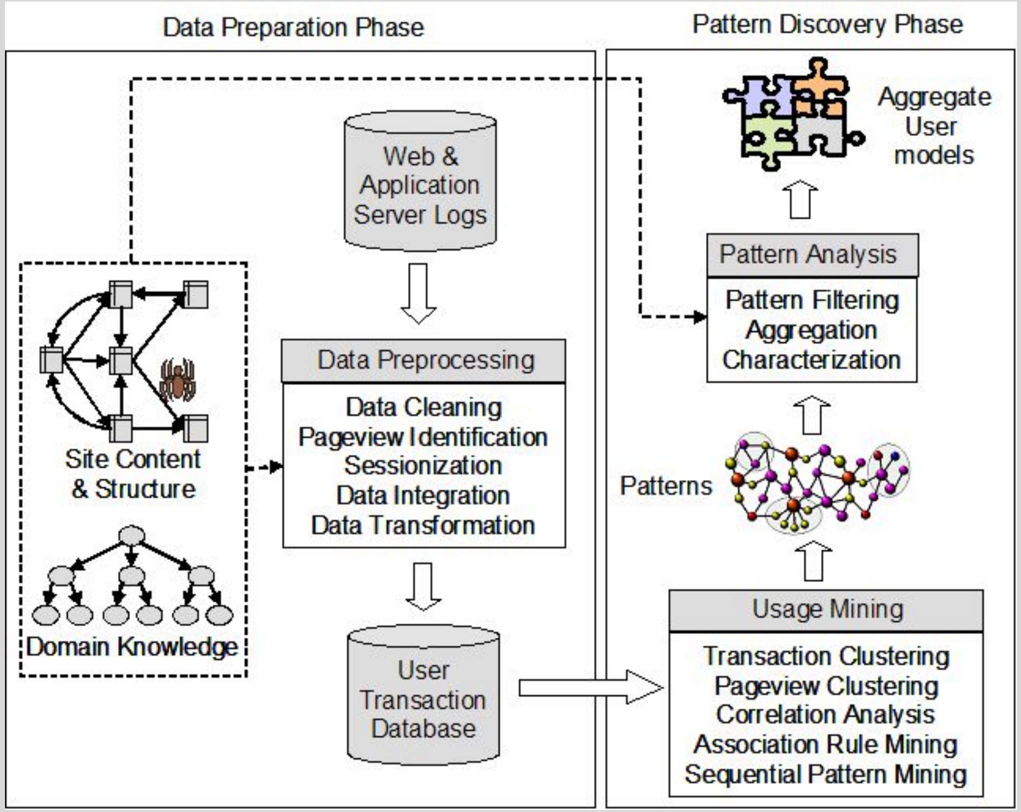
- Web mining is the application of data mining techniques to discover patterns from the Web.
- With the rapid growth of World Wide Web, Web mining becomes a very hot and popular topic in Web research.
- E-commerce and E-services are claimed to be the killer applications for Web mining, and Web mining now also plays an important role for E-commerce website and E-services to understand how their websites and services are used and to provide better services for their customers and users.

- **Applications of Web Mining:**
- ■E-commerce Customer Behavior Analysis
- ■E-commerce Transaction Analysis
- ■E-commerce Website Design
- ■E-banking
- ■M-commerce
- ■Web Advertisement
- ■Search Engine
- ■Online Auction.

- **Web mining can be divided into three different types, which are**
- ■Web usage mining,
- ■Web content mining
- ■Web structure mining.

- **a.Web Usage Mining**
- Web usage mining is the process of extracting useful information from server logs
- Web usage mining is the process of finding out what users are looking for on the Internet. Web Usage Mining focuses on techniques that could predict the behavior of users while they are interacting with the WWW.
- Web usage mining, discover user navigation patterns from web data, tries to discovery the useful information from the secondary data derived from the interactions of the users while surfing on the Web.
- Web usage mining collects the data from Web log records to discover user access patterns of web pages.

- There are several available research projects and commercial tools that analyze those patterns for different purposes.
- The insight knowledge could be utilized in personalization, system improvement, site modification, business intelligence and usage characterization.
- Some users might be looking at only textual data, whereas some others might be interested in multimedia data.
- Web Usage Mining is the application of data mining techniques to discover interesting usage patterns from Web data in order to understand and better serve the needs of Web-based applications.
- Usage data captures the identity or origin of Web users along with their browsing behavior at a Web site.


Data Preparation Phase
Pattern Discovery Phase
Web & Application Server Logs
Data Preprocessing
Data Cleaning
Pageview Identification
Sessionization
Data Integration
Data Transformation
Site Content & Structure
Domain Knowledge
User Transaction Database
Usage Mining
Transaction Clustering
Pageview Clustering
Correlation Analysis
Association Rule Mining
Sequential Pattern Mining
Patterns
Pattern Analysis
Pattern Filtering
Aggregation
Characterization
Aggregate User models

- **Classification of Web Usage mining**
- □**Web Server Data:**
- The user logs are collected by the Web server. Typical data includes IP address, page reference and access time.
- □**Application Server Data:**
- Commercial application servers have significant features to enable e-commerce applications to be built on top of them with little effort. A key feature is the ability to track various kinds of business events and log them in application server logs.
- □**Application Level Data:**
- New kinds of events can be defined in an application, and logging can be turned on for them thus generating histories of these specially defined events. It must be noted, however, that many end applications require a combination of one or more of the techniques applied in the categories above.

- **Pros & Cons of Web Usage Mining**
- Government agencies are using this technology to classify threats and fight against terrorism.
- The predicting capability of mining applications can benefit society by identifying criminal activities.
- Better customer relationship by giving them exactly what they need.
- attract and retain customers
- Cons: Invasion of Privacy

- **b.Web Content Mining**
- Web content mining targets the knowledge discovery, in which the main objects are the traditional collections of multimedia documents such as images, video, and audio, which are embedded in or linked to the web pages.
- It is also quite different from Data mining because Web data are mainly semi-structured and/or unstructured, while Data mining deals primarily with structured data.
- Web content mining is also different from Text mining because of the semi-structure nature of the Web, while Text mining focuses on unstructured texts.
- Web content mining thus requires creative applications of Data mining and / or Text mining techniques and also its own unique approaches.

- In the past few years, there was a rapid expansion of activities in the Web content mining area.
- This is not surprising because of the phenomenal growth of the Web contents and significant economic benefit of such mining.
- However, due to the heterogeneity and the lack of structure of Web data, automated discovery of targeted or unexpected knowledge information still present many challenging research problems.
- Web content mining could be differentiated from two points of view: Agent-based approach or Database approach.
- The first approach aims on improving the information finding and filtering.
- The second approach aims on modeling the data on the Web into more structured form in order to apply standard database querying mechanism and data mining applications to analyze it.

Web Content Mining Problems/Challenges :

- **·Data/Information Extraction:** Extraction of structured data from Web pages, such as products and search results is a difficult task. Extracting such data allows one to provide services. Two main types of techniques, machine learning and automatic extraction are used to solve this problem.
- **·Web Information Integration and Schema Matching:** Although the Web contains a huge amount of data, each web site (or even page) represents similar information differently. Identifying or matching semantically similar data is a very important problem with many practical applications.
- **·Opinion extraction from online sources:** There are many online opinion sources, e.g., customer reviews of products, forums, blogs and chat rooms. Mining opinions (especially consumer opinions) is of great importance for marketing intelligence and product benchmarking.

- ·**Knowledge synthesis:** Concept hierarchies or ontology are useful in many applications. However, generating them manually is very time consuming. A few existing methods that explores the information redundancy of the Web will be presented. The main application is to synthesize and organize the pieces of information on the Web to give the user a coherent picture of the topic domain.
- ·**Segmenting Web pages and detecting noise:** In many Web applications, one only wants the main content of the Web page without advertisements, navigation links, copyright notices. Automatically segmenting Web page to extract the main content of the pages is interesting problem.

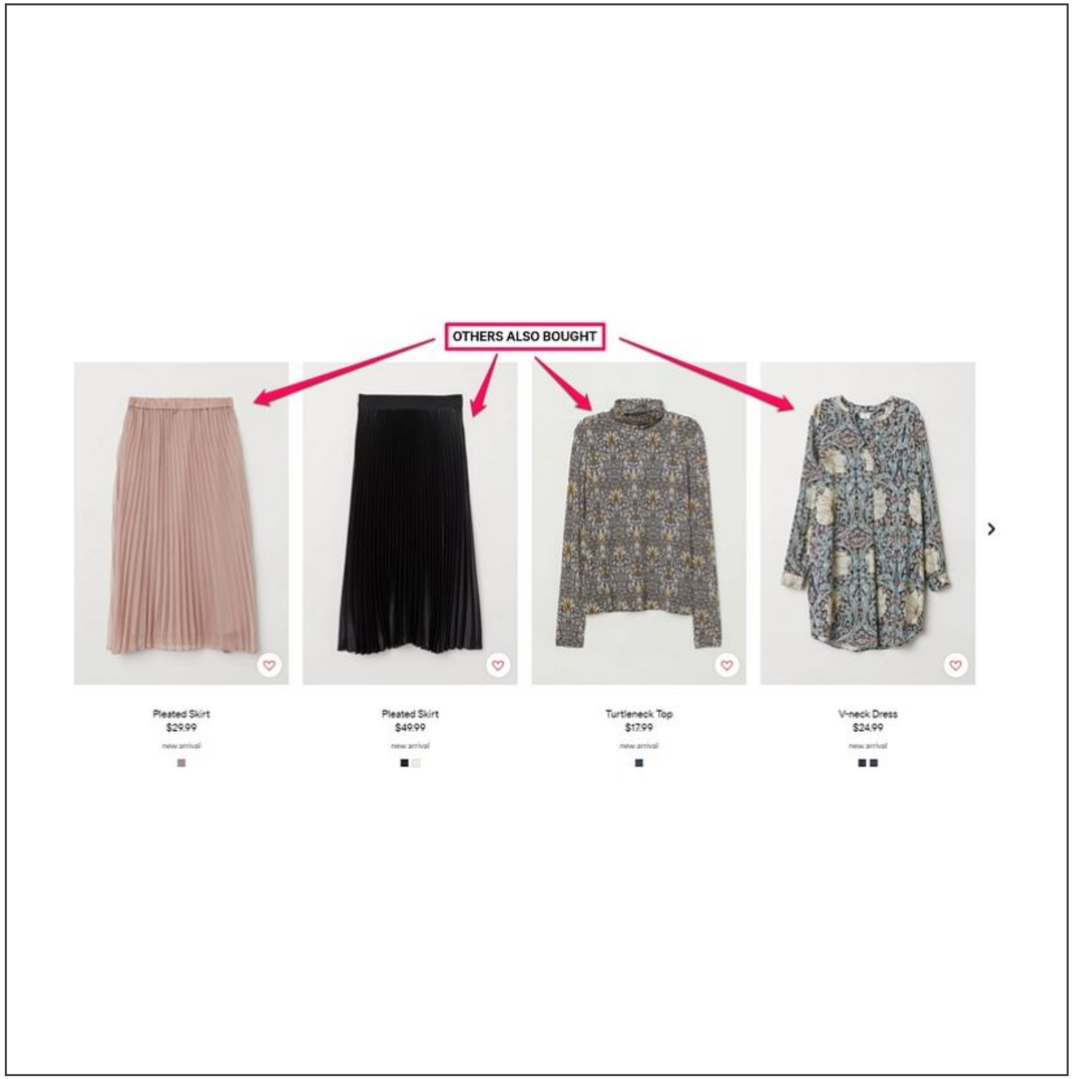
- **c.Web Structure Mining**
- Web structure mining focuses on analysis of the link structure of the web and one of its purposes is to identify more preferable documents.
- The different objects are linked in some way. The intuition is that a hyperlink from document A to document B implies that the author of document.
- A thinks document B contains worthwhile information. Web structure mining helps in discovering similarities between web sites or discovering important sites for a particular topic or discipline or in discovering web communities.
- Simply applying the traditional processes and assuming that the events are independent can lead to wrong conclusions. However, the appropriate handling of the links could lead to potential correlations, and then improve the predictive accuracy of the learned models.

- The goal of Web structure mining is to generate structural summary about the Web site and Web page.
- Technically, Web content mining mainly focuses on the structure of inner-document, while Web structure mining tries to discover the link structure of the hyperlinks at the inter-document level.
- Based on the topology of the hyperlinks, Web structure mining will categorize the Web pages and generate the information, such as the similarity and relationship between different Web sites.
- Web structure mining can also have another direction – discovering the structure of Web document itself.
- This type of structure mining can be used to reveal the structure (schema) of Web pages; this would be good for navigation purpose and make it possible to compare/integrate Web page schemes.
- This type of structure mining will facilitate introducing database techniques for accessing information in Web pages by providing a reference schema.

Recommender System

- A recommender system, or a recommendation system, is a subclass of information filtering system that seeks to predict the "rating" or "preference" a user would give to an item.
- They are primarily used in commercial applications.
- By drawing from huge data sets, the system's algorithm can pinpoint accurate user preferences.
- Once you know what your users like, you can recommend them new, relevant content.
- And that's true for everything from movies and music.
- Netflix, YouTube, Tinder, and Amazon are all examples of recommender systems in use. The systems entice users with relevant suggestions based on the choices they make.

- Recommender systems can also enhance experiences for:
- **News Websites**
- **Computer Games**
- **Knowledge Bases**
- **Social Media Platforms**
- **Stock Trading Support Systems**

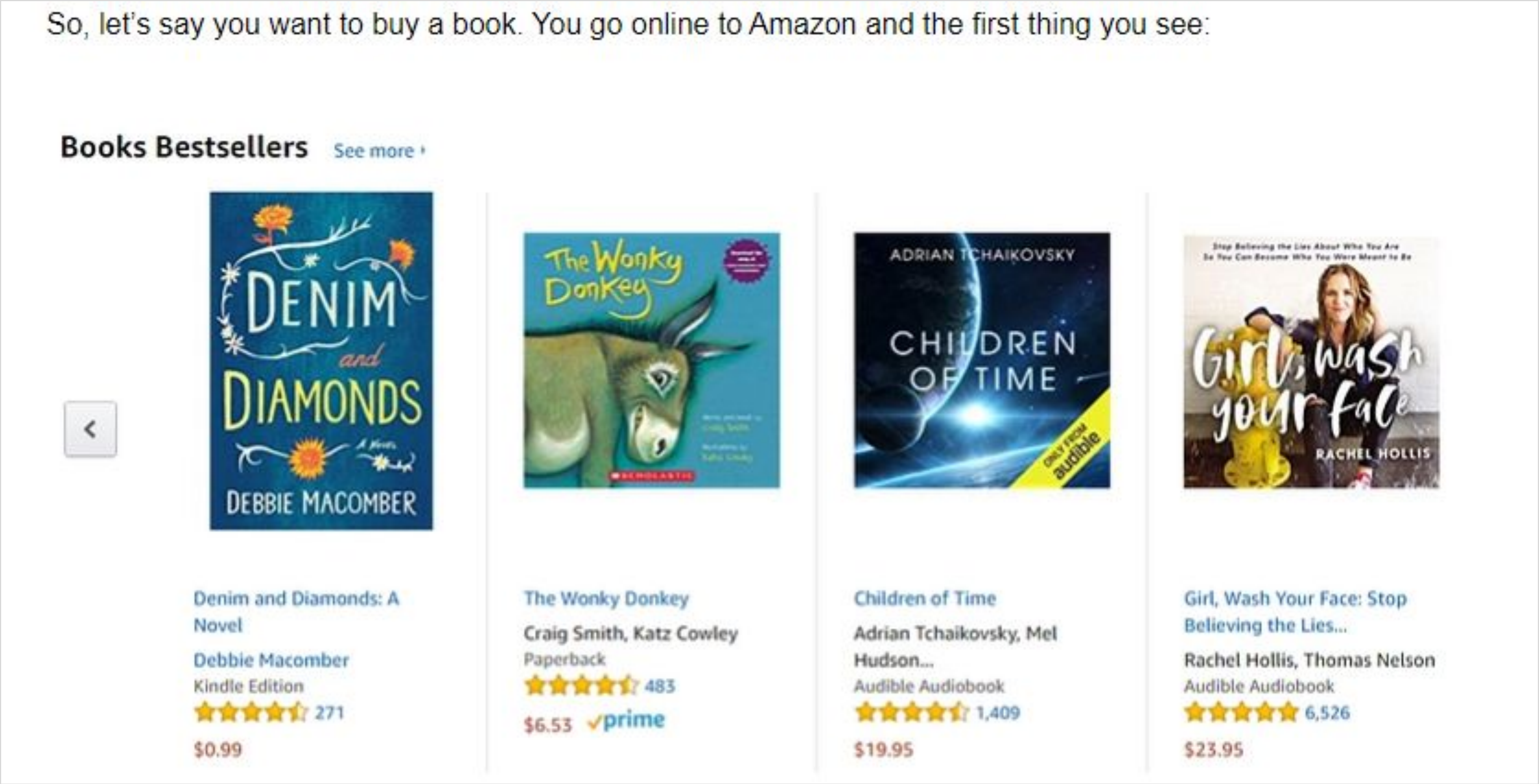


- Here’s an example of a recommender system in e-commerce.
- H&M served the following recommendations to users who clicked on “pleated skirt” as a potential buy:

So, what are the advantages of adding a recommender system to your website or software?

- Here’s a list of just a few:
- **Increase in sales thanks to personalized offers.**
- **Enhanced customer experience.**
- **Customer retention thanks to users feeling understood.**

- A recent study by Epsilon found that **90% of consumers find personalization appealing**.
- Plus, a further **80% claim they are more likely to do business with a company** when offered personalized experiences.
- The study also found that these **consumers are 10x more likely to become VIP customers**, who make more than 15 purchases per year.



So, let's say you want to buy a book. You go online to Amazon and the first thing you see:

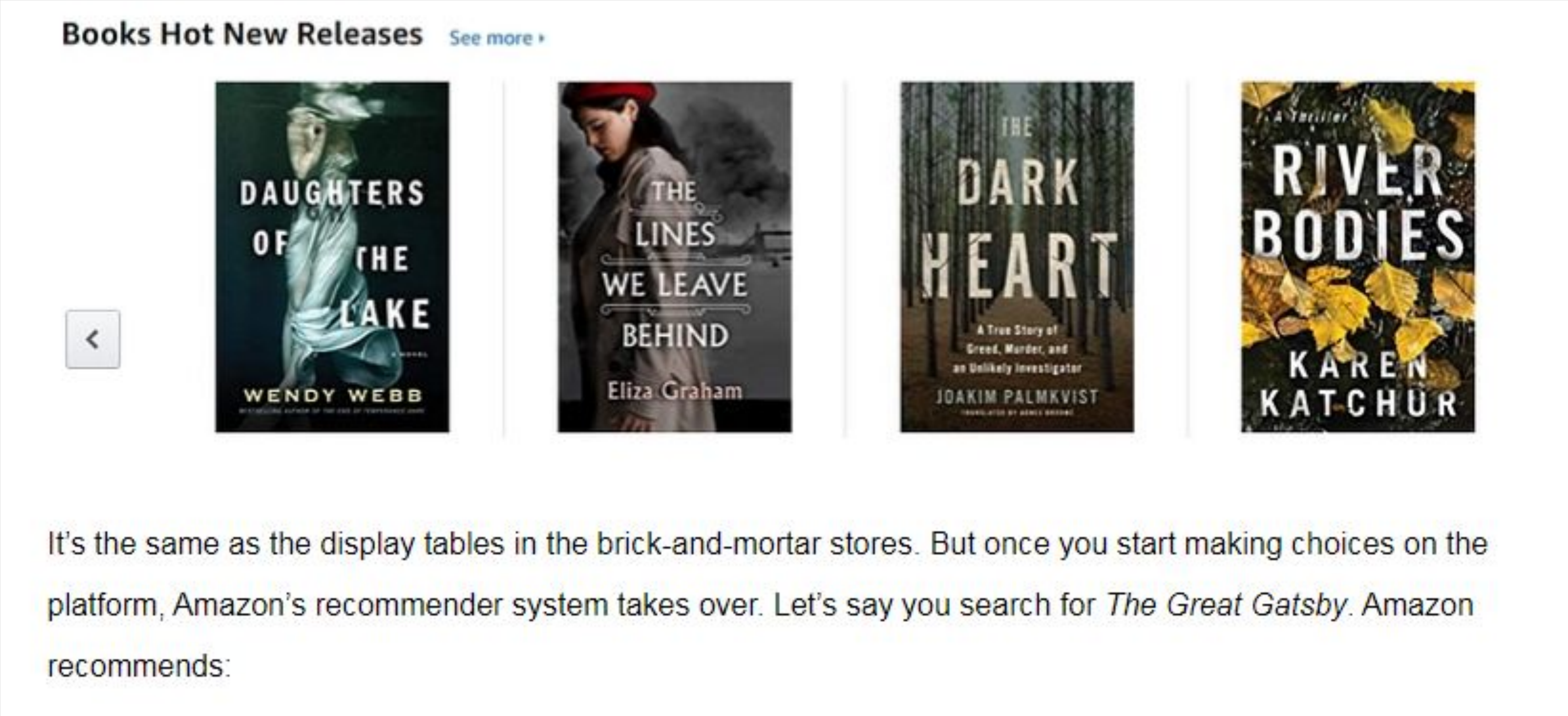
Books Bestsellers See more ›

Denim and Diamonds: A Novel
Debbie Macomber
Kindle Edition
★★★★☆ 271
$0.99

The Wonky Donkey
Craig Smith, Katz Cowley
Paperback
★★★★☆ 483
$6.53 ✓prime

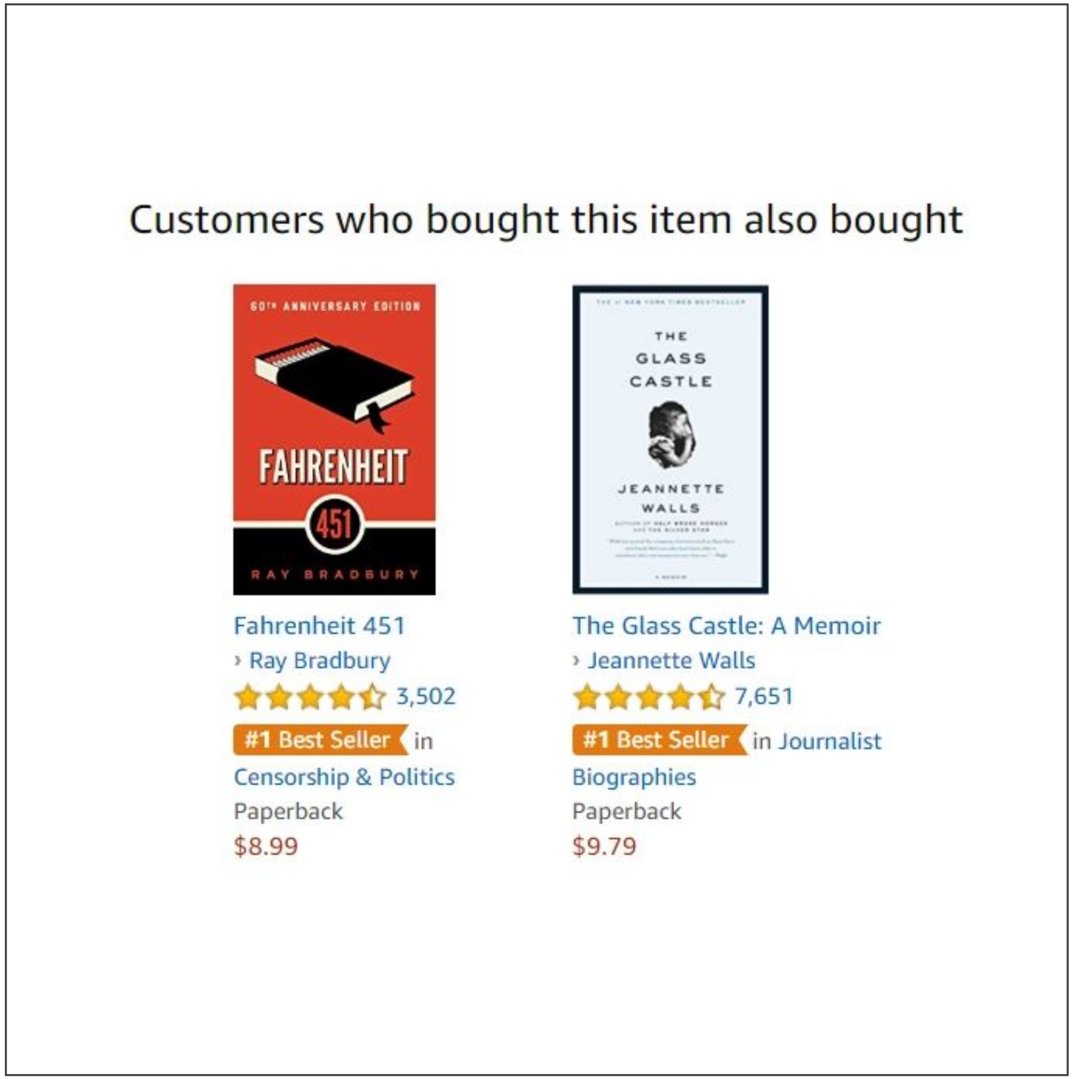
Children of Time
Adrian Tchaikovsky, Mel Hudson...
Audible Audiobook
★★★★☆ 1,409
$19.95

Girl, Wash Your Face: Stop Believing the Lies...
Rachel Hollis, Thomas Nelson
Audible Audiobook
★★★★★ 6,526
$23.95



Books Hot New Releases See more ›

It's the same as the display tables in the brick-and-mortar stores. But once you start making choices on the platform, Amazon's recommender system takes over. Let's say you search for *The Great Gatsby*. Amazon recommends:



Customers who bought this item also bought

Fahrenheit 451
› Ray Bradbury
★★★★☆ 3,502
#1 Best Seller in Censorship & Politics
Paperback
$8.99

The Glass Castle: A Memoir
› Jeannette Walls
★★★★☆ 7,651
#1 Best Seller in Journalist Biographies
Paperback
$9.79

◦Here the system served you *Fahrenheit 451*. That's because past Fitzgerald customers must have also bought Bradbury.

◦As an alternative, your recommender system could offer other Fitzgerald books.

Types of Recommender System

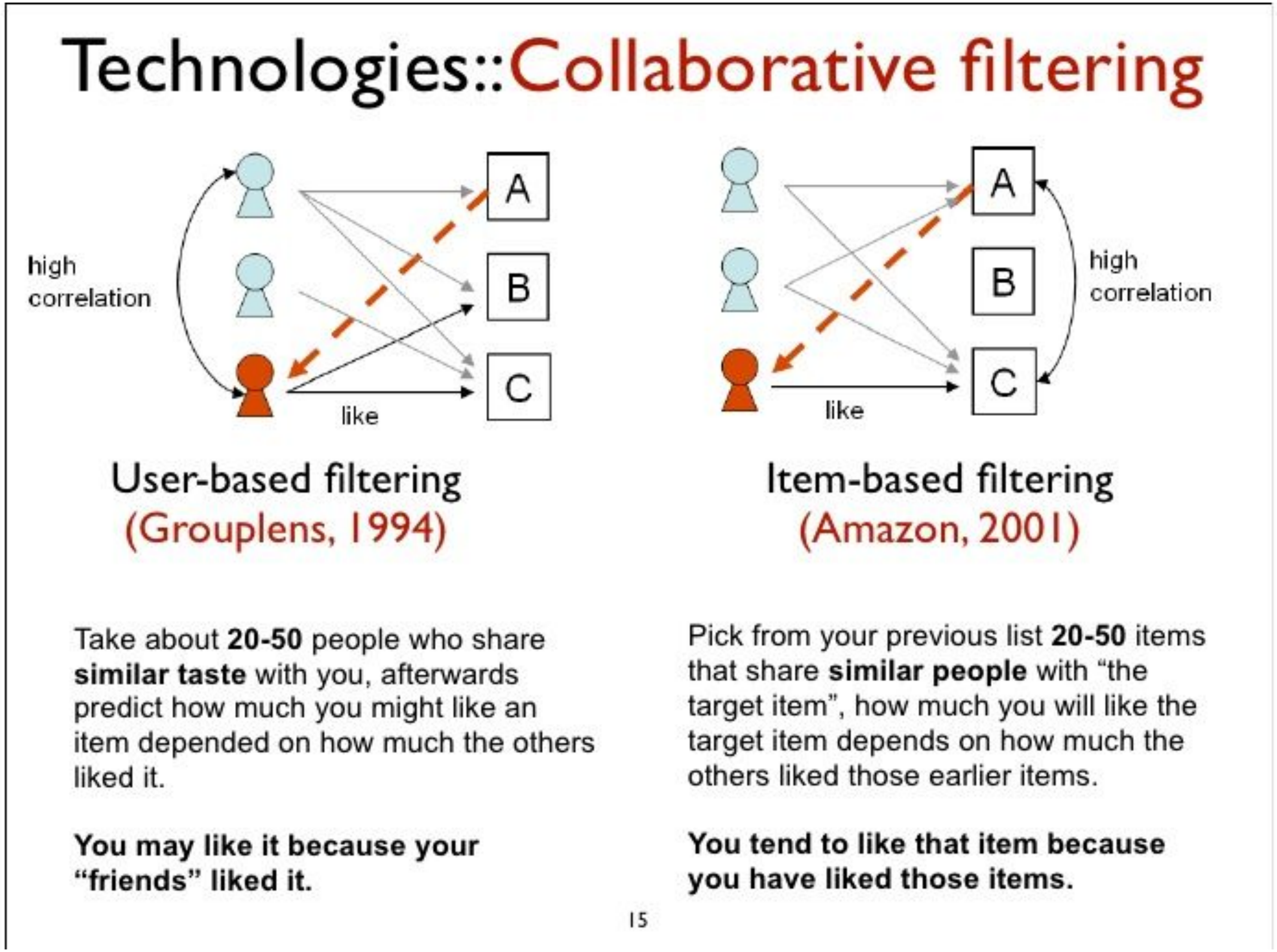
- Collaborative Filtering
- Content-based Filtering
- Hybrid (Combination of Both)

Collaborative Filtering Recommender Systems

- For starters, popular examples of collaborative filtering systems include Spotify, Netflix, and YouTube. But what does a collaborative filtering recommender system do?
- A collaborative filtering recommender system analyzes similarities between users and/or item interactions. Once the system identifies similarities, it serves users recommendations. In general, users see items that similar users liked.
- There are different types of collaborative filtering systems including:
- **Item-item Collaborative Filtering**
- **User-user Collaborative Filtering**

Item-item Collaborative Filtering

- **Item**-**item collaborative filtering**, or **item**-based, or **item-to-item**, is a form of **collaborative filtering** for recommender systems based on the similarity between items calculated using people's ratings of those items.
- **Item**-**item collaborative filtering** was invented and used by Amazon.com in 1998.
- An item-item filtering algorithm analyzes product associations taken from user ratings. Users then see recommendations based on how they rate individual products.
- For example, you rate a book or movie as a 10/10. Now, you will see the top-rated books or movies with similar attributes. Below is an example from Goodreads.
- I created a special list for booaks that I gave five-star ratings. Goodreads then recommends me the highest-ranked books from similar readers' lists.



Technologies::Collaborative filtering

User-based filtering
(Grouplens, 1994)

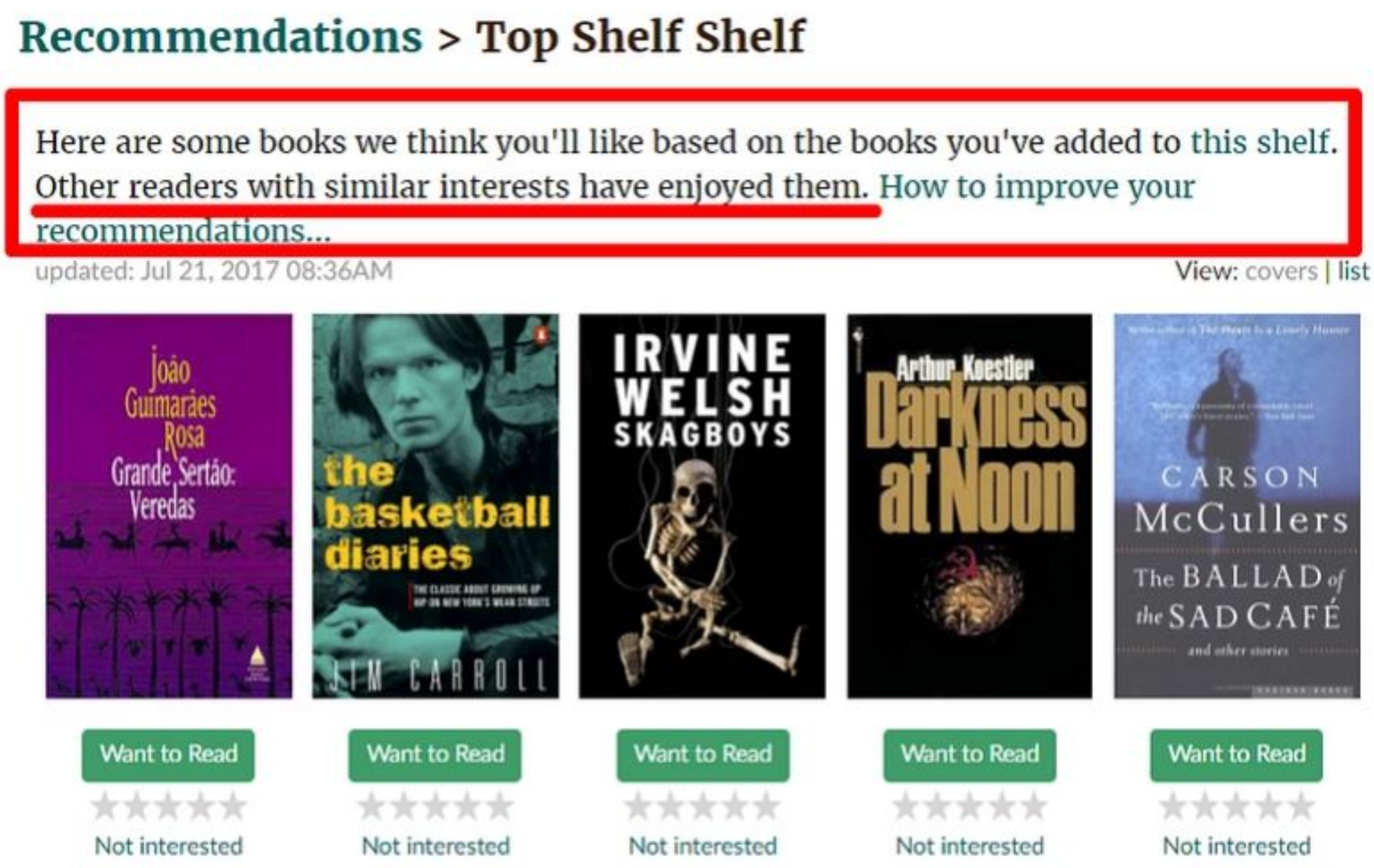
Item-based filtering
(Amazon, 2001)

Take about **20-50** people who share **similar taste** with you, afterwards predict how much you might like an item depended on how much the others liked it.

You may like it because your "friends" liked it.

Pick from your previous list **20-50** items that share **similar people** with "the target item", how much you will like the target item depends on how much the others liked those earlier items.

You tend to like that item because you have liked those items.

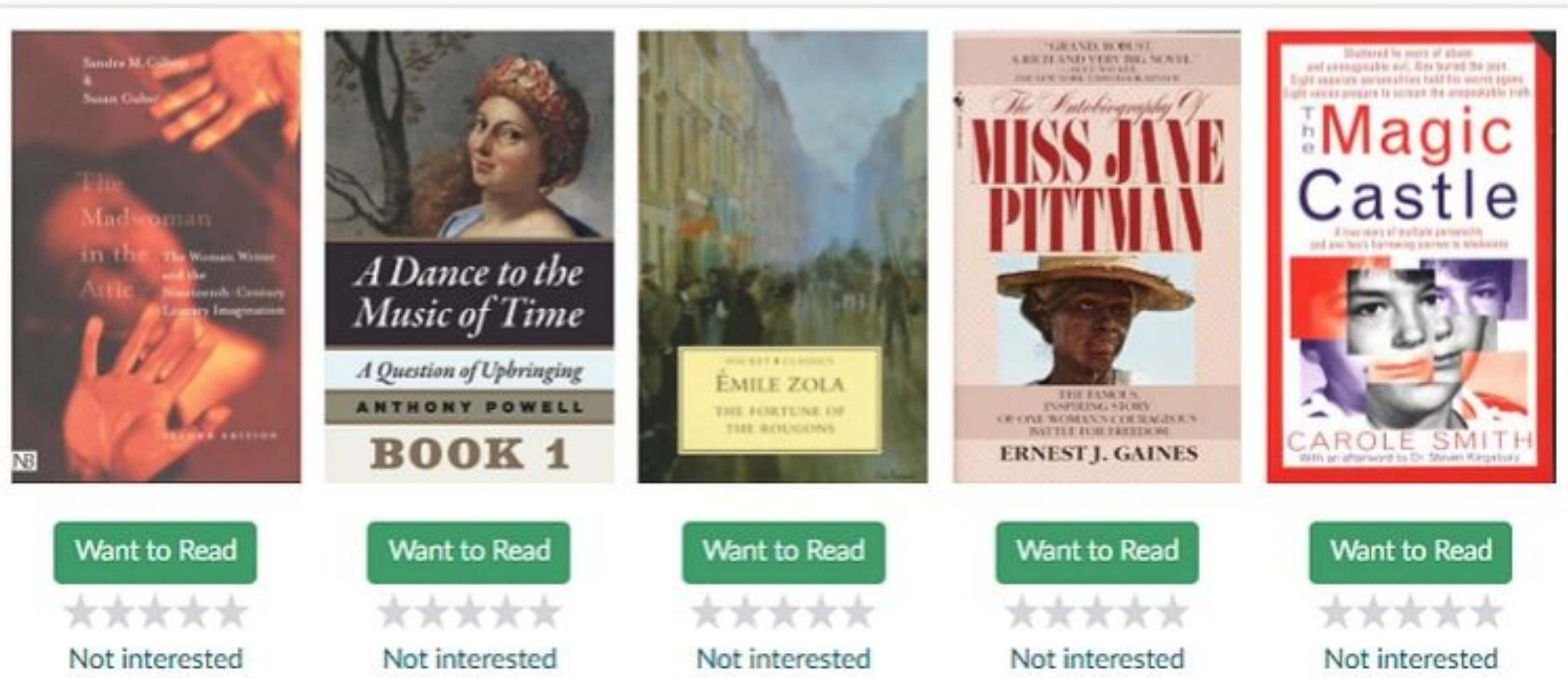


Recommendations > Top Shelf Shelf

Here are some books we think you'll like based on the books you've added to this shelf. Other readers with similar interests have enjoyed them. How to improve your recommendations...

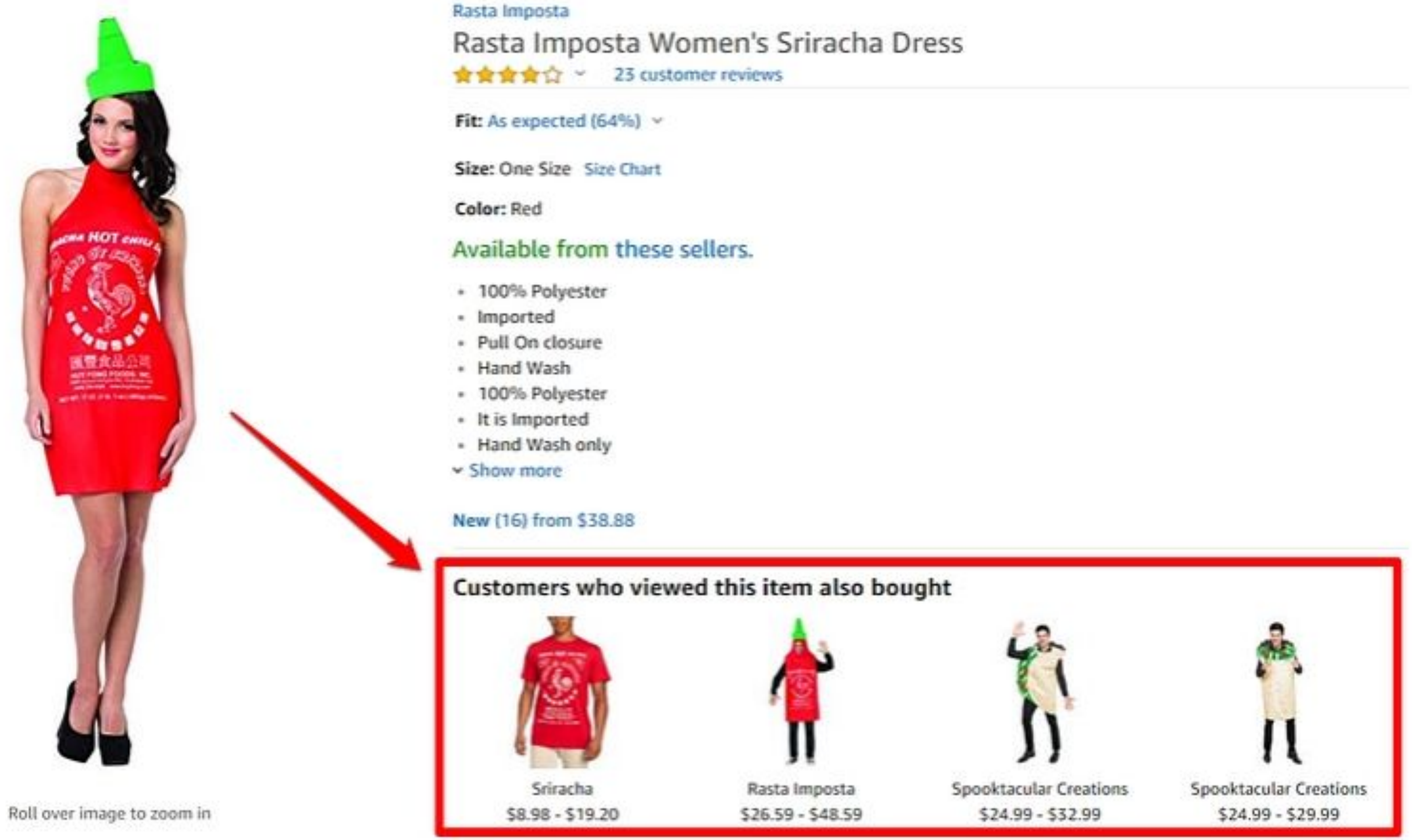
updated: Jul 21, 2017 08:36AM

View: covers | list



It's not always easy to get users to give items ratings. That's why item-item filtering can be as simple as clicking on a dress and seeing more dresses.

- Ever come across the *“people who viewed this item also bought”* copy under a product?
- That’s right. That’s also an item-item filtering system.
- Amazon invented item-item filtering for their recommender system. Item filtering works best when you have more users on your platform than items.



Rasta Imposta

Rasta Imposta Women's Sriracha Dress

★★★★☆ 23 customer reviews

Fit: As expected (64%)

Size: One Size Size Chart

Color: Red

Available from these sellers.

- 100% Polyester
- Imported
- Pull On closure
- Hand Wash
- 100% Polyester
- It is Imported
- Hand Wash only

Show more

New (16) from $38.88

Customers who viewed this item also bought

Sriracha	Rasta Imposta	Spooktacular Creations	Spooktacular Creations
$8.98 - $19.20	$26.59 - $48.59	$24.99 - $32.99	$24.99 - $29.99

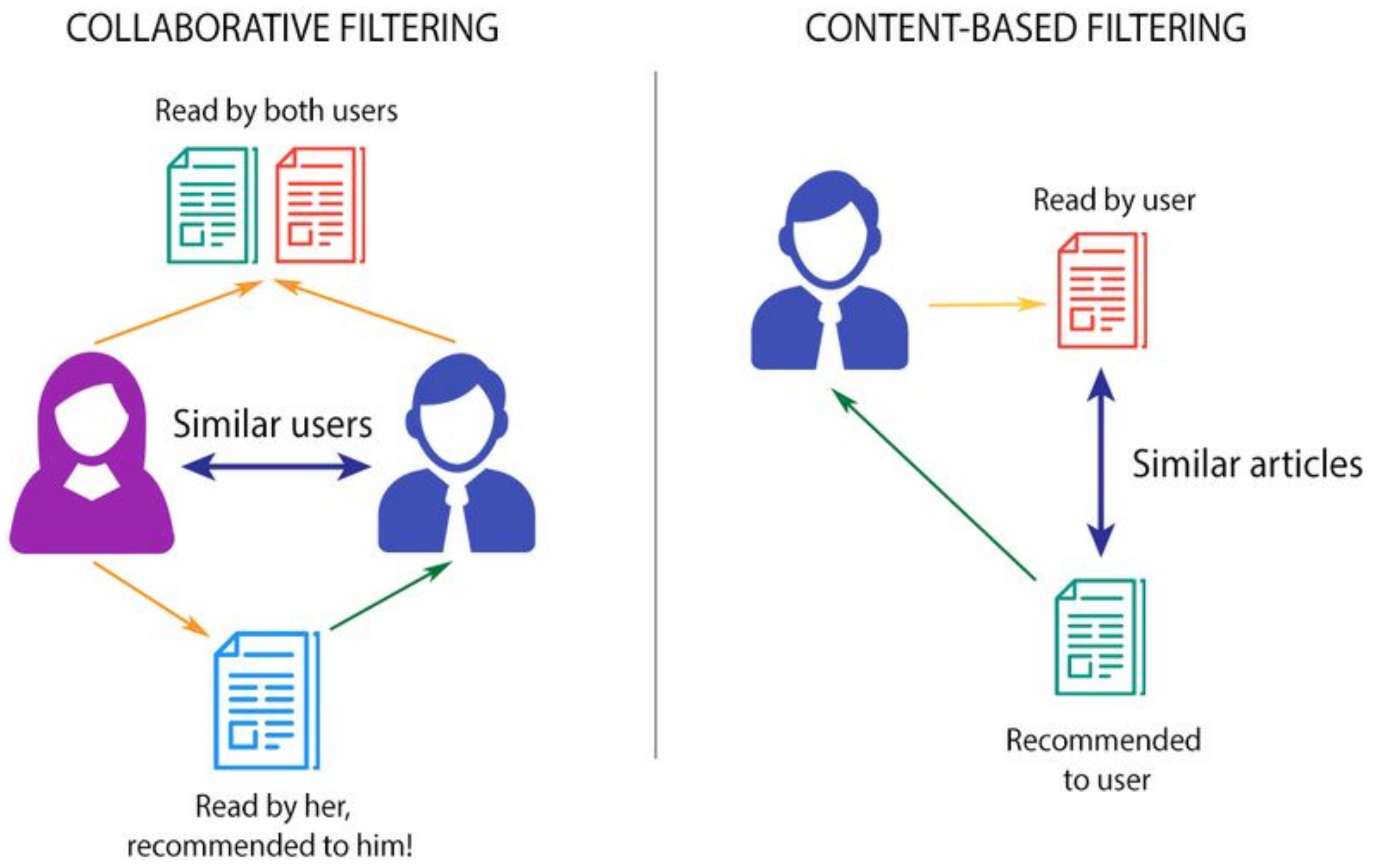
Roll over image to zoom in

User-user Collaborative Filtering

- The other kind of collaborative filtering takes the similarity of user tastes into consideration.
- So, user-user collaborative filtering doesn’t serve you items with the best ratings.
- Instead, you join a cluster of other people with similar tastes and you see content based on historic choices.
- Let’s say you use YouTube for the first time. You play a Beyonce song. The system clusters you with other users who also like Beyonce. Then the YouTube recommendation system shows you other videos chosen by users in your cluster. The more choices you make, the more relevant the results.

Content Based Recommender Systems

- Content based filtering uses the characteristics or properties of an item to serve recommendations. Characteristic information includes:
- **Characteristics of Items (Keywords and Attributes)**
- **Characteristics of Users (Profile Information)**
- Let's use a movie recommendation system as an example. Characteristics for the item *Harry Potter and the Sorcerer's Stone* might include:
- **Director Name** – Chris Columbus
- **Genres** – Adventure, Fantasy, Family (IMDB)
- **Stars** – Daniel Radcliffe, Rupert Grint, Emma Watson
-


COLLABORATIVE FILTERING
Read by both users
Similar users
Read by her,
recommended to him!
CONTENT-BASED FILTERING
Read by user
Similar articles
Recommended
to user

- A content based recommender works with data that the user provides, either explicitly (rating) or implicitly (clicking on a link).
- Based on that data, a user profile is generated, which is then used to make suggestions to the user.
- As the user provides more inputs or takes actions on the recommendations, the engine becomes more and more accurate.

Pitfalls of Different Types of Recommender Systems

- **Types of Recommender Systems Problems – The Collaborative Filtering Problem**
- Collaborative filtering needs a lot of data to create relevant suggestions. So, when you start using a platform with a collaborative filtering system, you start cold.
- The cold start problem in recommender systems is common for collaborative filtering systems.
- For example, when John visits YouTube for the first time, the system has to wait for him to watch several videos. Only then can it serve him relevant recommendations for other videos.

- **Cold Start Problem**
- □Cold start is a potential problem in computer-based information systems which involve a degree of automated data modeling.
- □It concerns the issue that the system cannot draw any inferences for users or items about which it has not yet gathered sufficient information.
- □In the collaborative filtering approach, the recommender system would identify users who share the same preferences (e.g. rating patterns) with the active user, and propose items which the like-minded users favored (and the active user has not yet seen). Due to the cold start problem, this approach would fail to consider items which no-one in the community has rated previously.
- **Solution**
- ■In scenarios involving interface agents, the cold start problem may be overcome by introducing an element of collaboration amongst agents assisting various users (Community based recommendation).
- ■This way, novel situations may be handled by requesting other agents to share what they have already learnt from their respective users

- **Types of Recommender Systems Solutions – The Collaborative Filtering Solution**
- A solution to the cold start problem in recommender systems is clustering data with attribute similarities. Let’s go back to our YouTube example.
- John visits YouTube for the first time. The first video he selects is a Beyonce video. As mentioned before, the platform will cluster John with other users who watched the same video.
- It could also add him to other clusters. Let’s say the video belongs to the “pop song” cluster. Needless to say, the pop song cluster is populated with pop songs,
- Now, the system can recommend other songs based on the following criteria:
- Other Beyonce Listeners’ Choices
- Other Songs in the Pop Cluster

- **Types of Recommender Systems Problems – The Content Based Filtering Problem**
- The problem with content-based recommender systems is that they are restrictive. You click on a dress and you see more dresses. The system is incapable of knowing that your interests go beyond liking dresses.
- **Types of Recommender Systems Solutions – The Content Based Filtering Solution**
- Again, a common solution is to ask users upfront about what kind of things they like. And as users interact with your site, you can use historical data to recommend them more tailored choices.
- The customer buys a dress and some shoes. Now, you know that she likes both.

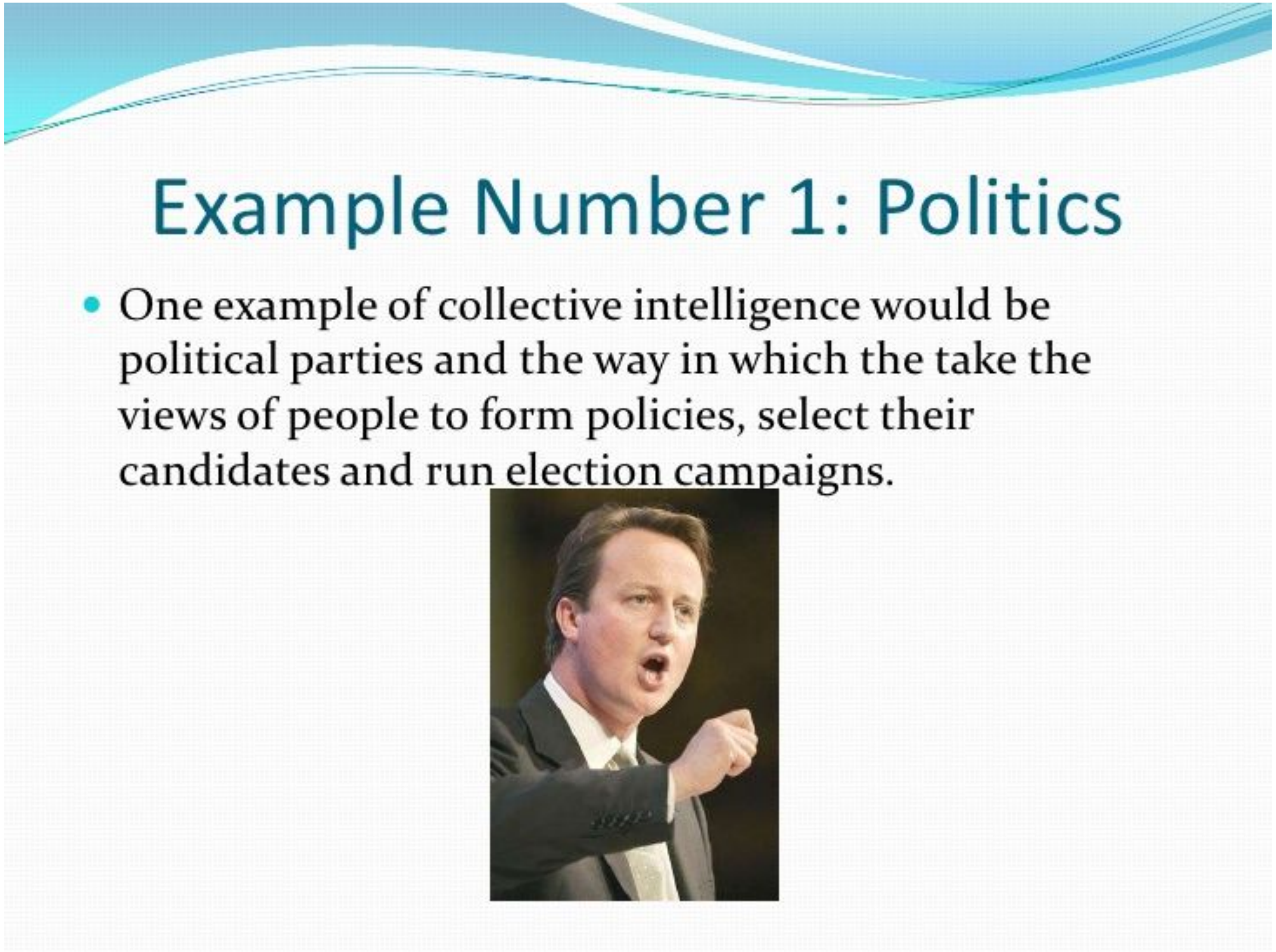
7.8Collective Intelligence

- It is a shared or group intelligence that emerges from the collaboration, collective efforts, and competition of many individuals. The concept is used in sociology, business, computer science and mass communications.
- "*It is a form of universally distributed intelligence, constantly enhanced, coordinated in real time, and resulting in the effective mobilization of skills*."
- Collective intelligence is a mass collaboration. In order for this concept to happen, four principles need to exist:

- ■**Openness:** Sharing ideas and intellectual property: By allowing others to share ideas, it gains significant improvement.
- ■**Peering:** Peering where users are free to modify and develop programs, provided that they make it available for others.
- ■**Sharing:** Companies have started to share some ideas while maintaining some degree of control over others.
- ■**Acting Globally:** The internet is widespread, therefore a company has no geographical boundaries and may access new markets, ideas and technology.

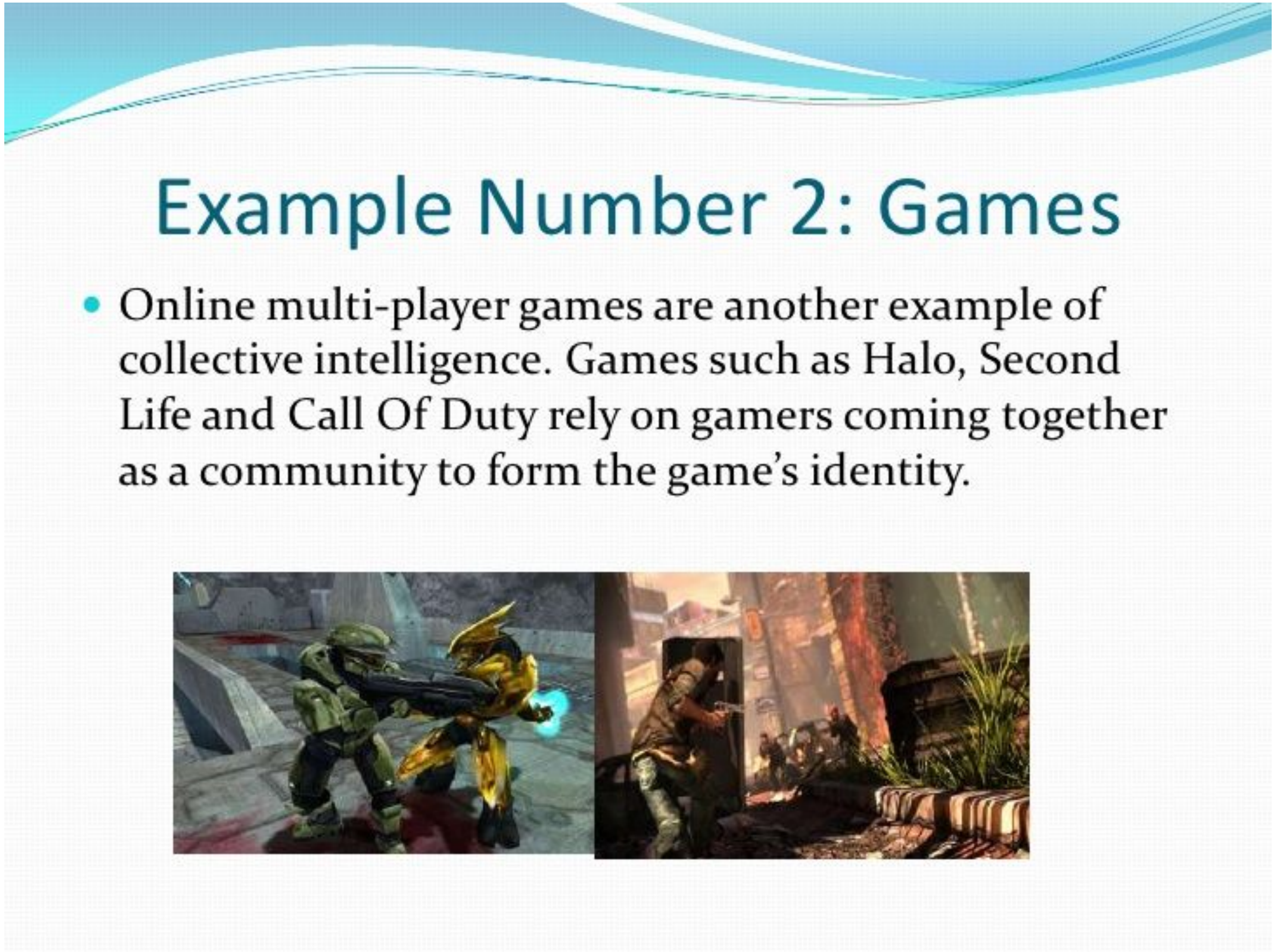
- Collective intelligence can be harnessed from social media through a variety of means and can be beneficial to your organization.
- Surveys and polls are available on sites like Facebook and LinkedIn, allowing you to identify trends and patterns in people's opinions.
- By monitoring the Likes, shares, and comments on social networking sites, you will eventually see certain patterns arise in people's viewpoints that show the popularity of one opinion over another.
- To give an example, let's say that you owned a shoe company and wanted to identify which product line would sell the most.

- By uploading photos to a social bookmarking site , people can click a link indicating they like the product .
- These photos can be shared with others on other social networking sites like Facebook, increasing exposure to the product.
- By monitoring the reactions of people, you'll see trends where a majority of people liked one shoe over another, and thereby predict that it will sell better than others.
- While the results might not have a guarantee, they would tend to be more accurate than the opinion of a single or small group of decision makers.



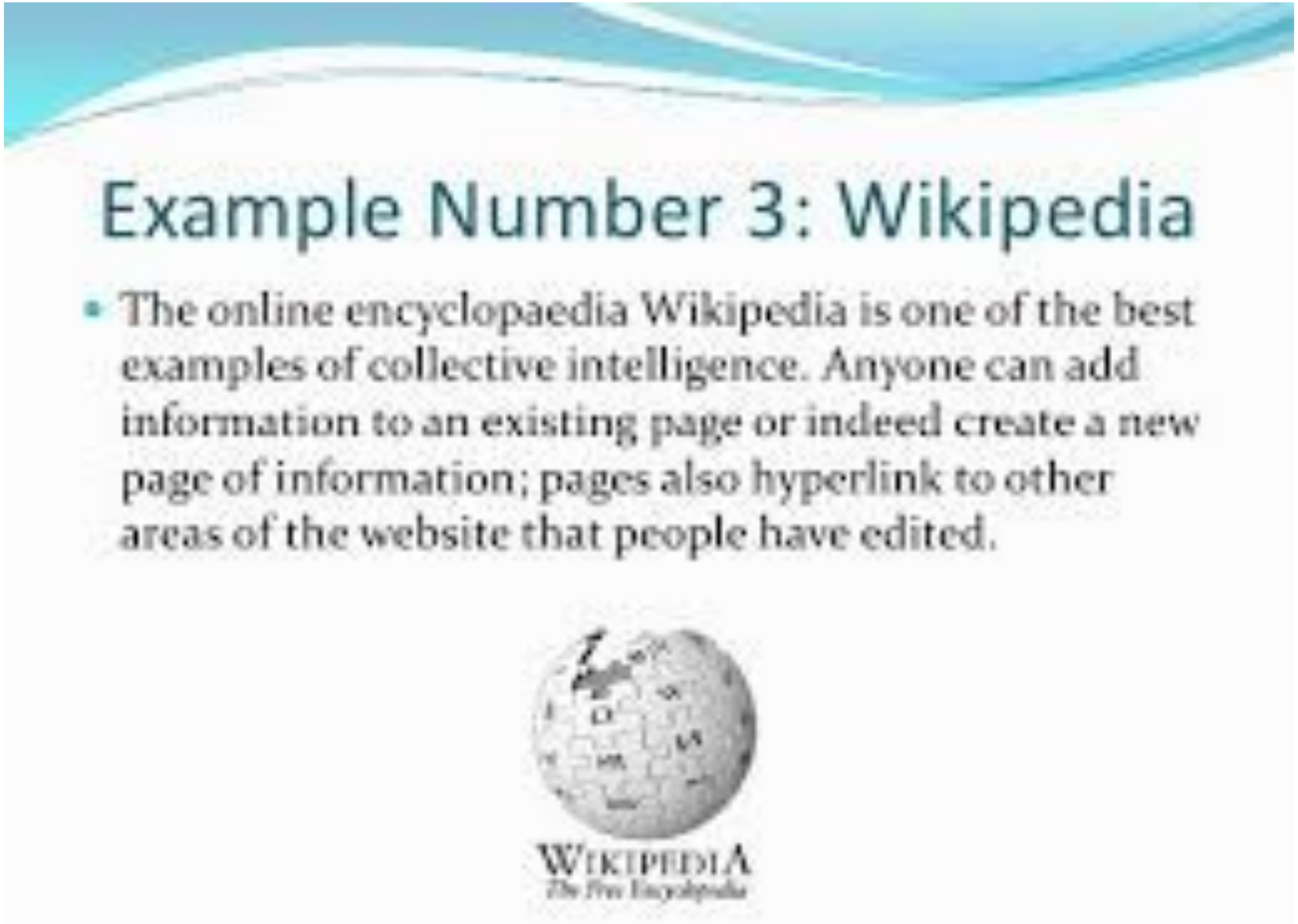
Example Number 1: Politics

- One example of collective intelligence would be political parties and the way in which the take the views of people to form policies, select their candidates and run election campaigns.



Example Number 2: Games

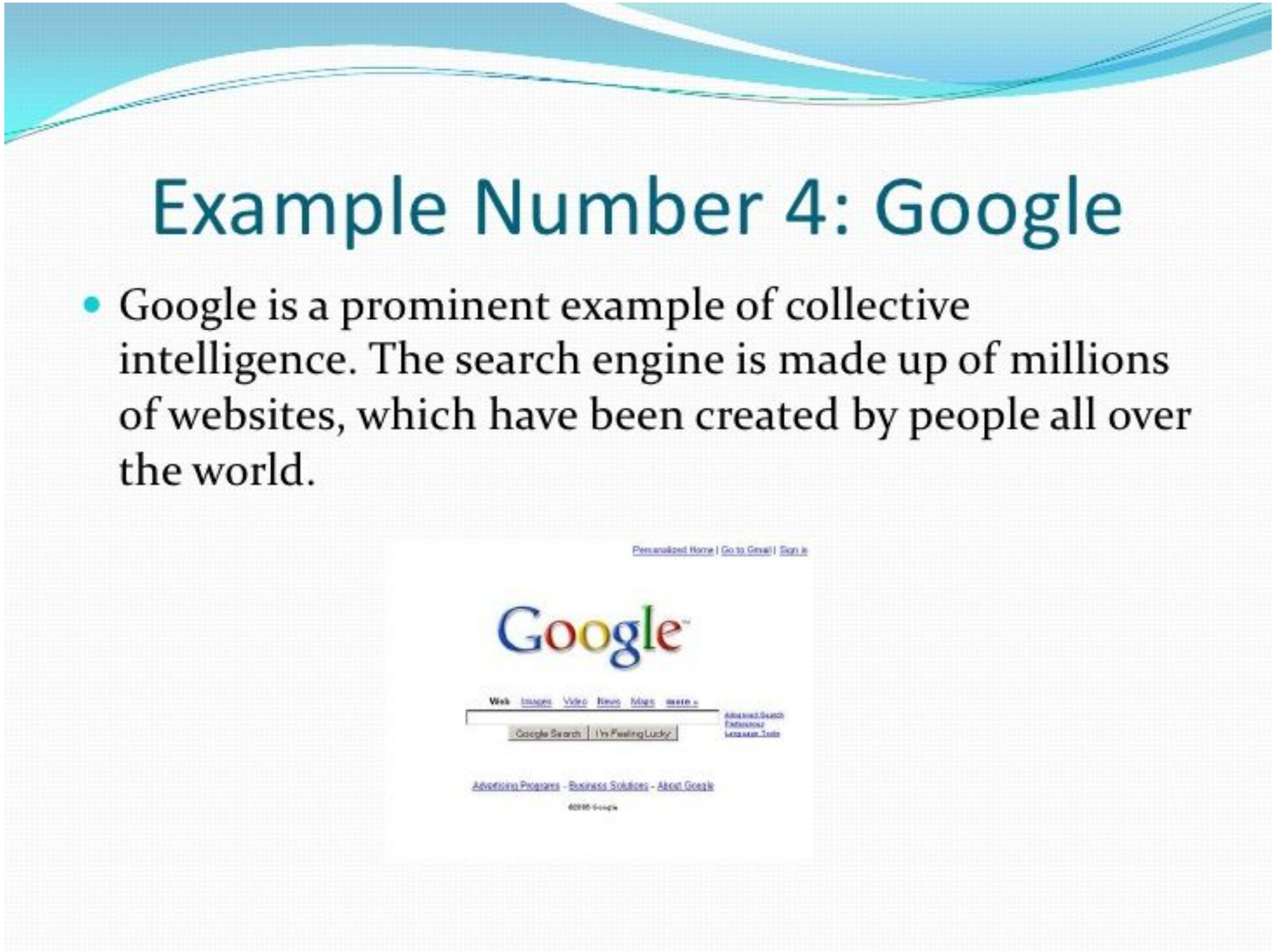
- Online multi-player games are another example of collective intelligence. Games such as Halo, Second Life and Call Of Duty rely on gamers coming together as a community to form the game's identity.



Example Number 3: Wikipedia

- The online encyclopaedia Wikipedia is one of the best examples of collective intelligence. Anyone can add information to an existing page or indeed create a new page of information; pages also hyperlink to other areas of the website that people have edited.

WIKIPEDIA
The Free Encyclopedia



Example Number 4: Google

- Google is a prominent example of collective intelligence. The search engine is made up of millions of websites, which have been created by people all over the world.

Personalized Home | Go to Gmail | Sign in
Google
Web Images Video News Maps more »
Advanced Search
Preferences
Language Tools
Google Search
I'm Feeling Lucky
Advertising Programs - Business Solutions - About Google
©2006 Google

Final example: Amazon

- If a person has a Amazon account they can buy or sell products to other people with accounts this is collective intelligence because the people are making up the website.
- The website also recommends items that may also interest you judging on what you have already looked at which is collective intelligence also.
- Things such as customer reviews can also be heavily influential when choosing a product. You are essentially basing your opinion off of the opinions of other members of the public.

- https://www.iteratorshq.com/blog/an-introduction-recommender-systems-9-easy-examples/
- https://towardsdatascience.com/various-implementations-of-collaborative-filtering-100385c6dfe0