CS6120:
Intelligent Media Systems

Dr. Derek Bridge
School of Computer Science & Information Technology
UCC

Lecturer

- Dr. Derek Bridge
  – Room G61, WGB
  – d.bridge@cs.ucc.ie
  – www.cs.ucc.ie/dbridge.html

Teaching & Examining

- Lectures:
  – 16:00-18:00 Tuesdays, WGB G:18
- Labs:
  – 14:00-15:00 Thursdays, WGB 1:15
- Handouts:
  – copies of slides from my web site
- Written exam:
  – 90 mins in the Summer, 60 marks
- Continuous assessment:
  – Python programming project, 40 marks

Outline (tentative)

1. Introduction
2. User Identification
3. Web Search
4. Non-Personalized Recommendation
5. Content-Based Recommendation
6. Collaborative Recommendation
7. Evaluation of Recommenders
8. Item-Based and Hybrid Recommendation
9. Knowledge-Based Recommendation
10. Social Recommendation
11. TBA
12. TBA

INTELLIGENT MEDIA SYSTEMS

INTELLIGENT   WEB   SYSTEMS

History of the Web 1

Static web pages → Dynamic web pages → Interactive web pages

Static Web Pages

request for a web page

client computer running a browser

response – a copy of the web page

web server computer

Dynamic Web Pages

request to execute a program

program executes here

client computer running a browser

response – the output from the program

web server computer

Interactive Web Pages

program
executes
here

request
for a program

client computer
running a browser

response –
a copy of the program

web server computer

History of the Web 2

- From
  - core technologies (HTTP, URL, HTML)
- Through
  - several versions (e.g. of HTML, CSS)
- To
  - the current rich ecosystem

See www.evolutionoftheweb.com

History of the Web 3

- Client-side devices

desktop PCs

laptops

mobile
phones

smart
phones &
tablets

embedded
& wearable
sensors

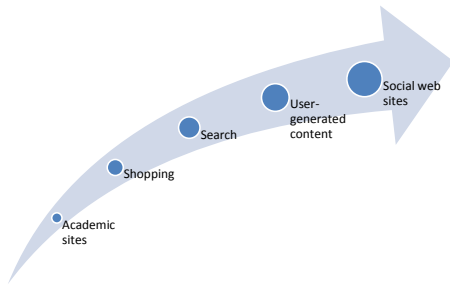## History of the Web 4

- Social web sites
- User-generated content
- Search
- Shopping
- Academic sites

Academic Web Sites

- Tim Berners-Lee
  – his "Vague, but exciting" proposal

- Early web sites
  – Academics and their institutions
  – Public bodies
  – A few businesses

Shopping

- …in the material world
  – Physical limitations on range of items available
  – Irreproducibality or limits on reproducibility of items
  – Can experience the item (test drive it, try it on, try it out, taste it,…)
  – Guidance from salesperson or other intermediary (librarian, travel agent, tutor,…)

- …online
  – Lower physical limitations means wider choice
  – Some items (digital ones) are infinitely reproducible
  – Cannot experience physical items; but may be able to play/look at samples of digital items
  – No human intermediary ('disintermediation')

The Long Tail

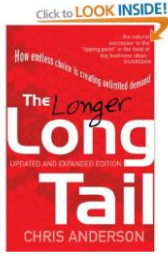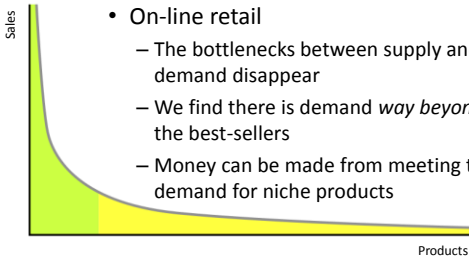## The Long Tail

- On-line retail
  - The bottlenecks between supply and demand disappear
  - We find there is demand *way beyond* the best-sellers
  - Money can be made from meeting this demand for niche products

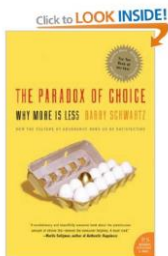## The Paradox of Choice

The Paradox of Choice

- Overwhelmed by choice
  - Choice no longer liberates, but debilitates – even tyrannizes!
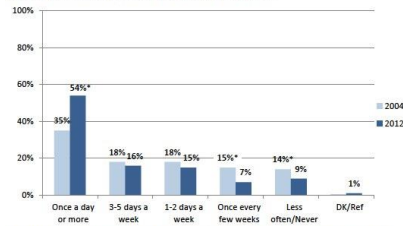 - Leads to bad decisions, anxiety, stress, dissatisfaction – even depression
- Really?

Search

Search users are turning to search engines more frequently
% of adult search users who use a search engine to find information....

Source: The Pew Research Center's Internet & American Life Project Winter 2012 Tracking Survey, January 20-February 19, 2012. N=2,253 adults, age 18 and older, including 901 cell phone interviews. Interviews conducted in English and Spanish. An asterisk (*) indicates a significant difference across years at the .95 confidence level.

User-Generated Content (UGC)

Web 2.0/ The Read-Write Web

- Users no longer just passively consume web content
  - They now contribute content too

Examples

- Upload
  - photos (Flickr, Picassa,...)
  - videos (YouTube, vimeo,..)
  - Blog posts (WordPress, Blogger, Tumblr,...)
- Edit articles (e.g. Wikipedia & other wikis)
- Express opinion by
  - rating, voting (inc. up-votes, down-votes),...
  - writing reviews
  - posting comments
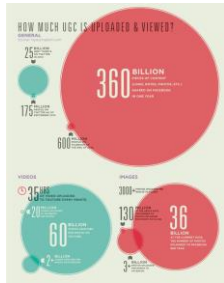- Discuss on forums, bulletin boards & Q&A sites (e.g. boards.ie)

User-Generated Content

http://www.fastcodesign.com/1664377/infographic-of-the-day-the-alchemy-behind-facebook-and-youtube, 2011

Social Web Sites

Social networking sites
- E.g. Facebook, LinkedIn
- Users
  - construct personal profiles
  - connect with other users
  - exchange status updates with their connections

Social media sites
- E.g. YouTube, Instagram
- Users
  - upload UGC
  - connect with other users
  - share the UGC with their connections

Social Web Sites

- New medium for communication
- New means of building communities
- New source of information
- New form of entertainment
- New platform for self-expression

Information Overload
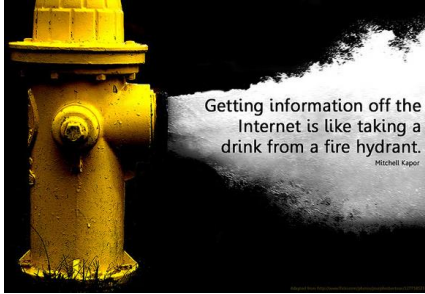

Getting information off the Internet is like taking a drink from a fire hydrant.
Mitchell Kapor

History of the Web 5

- Content discovery

Browsing → Searching → Sharing → Recommendation

INTELLIGENT WEB SYSTEMS

ADAPTIVE   WEB SYSTEMS

Adaptive Systems

- Automatically tailor content, presentation and behaviour to the user and the context of use
- Two important special cases:
  - *personalization*
    - adapt to the user's interests, tastes, preferences, short- and long-term goals, skills & knowledge,…
  - *contextualization*
    - adapt to time, place, user's companions,…
- Not *customization*, where the user configures the system

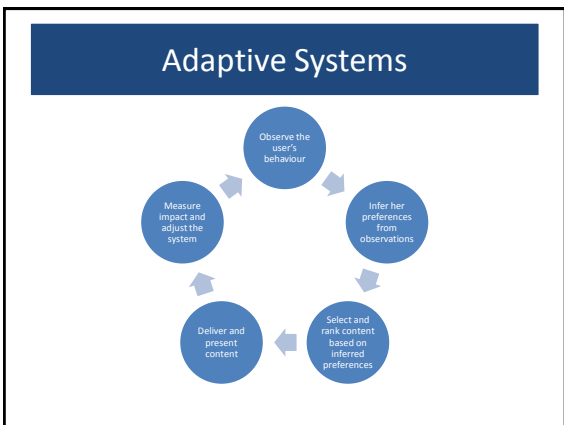## Adaptive Systems

Observe the user's behaviour → Infer her preferences from observations → Select and rank content based on inferred preferences → Deliver and present content → Measure impact and adjust the system

Observations
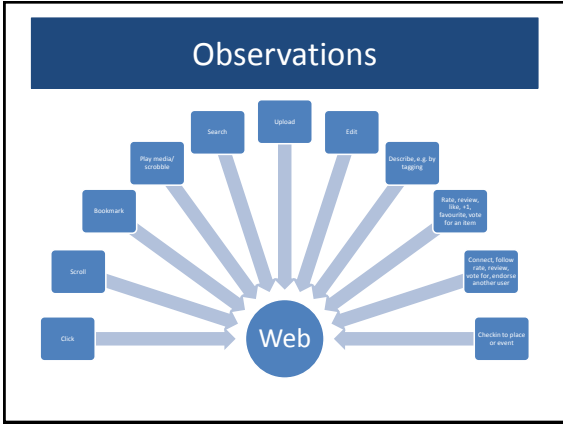
Case Study

N. Negroponte (1995): *Being Digital,* Vintage Books

Case Study
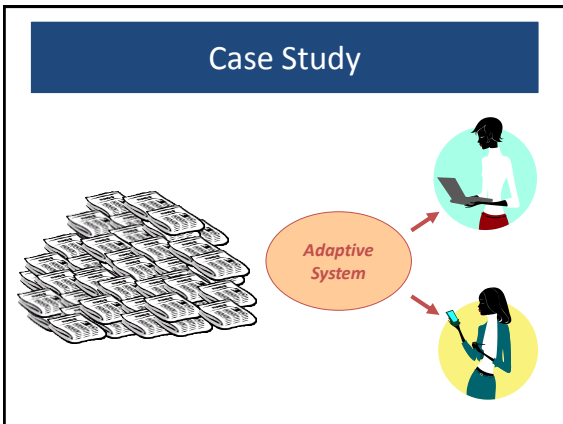
Google News

- Google News is a news aggregator
  - developed by Krishna Bharat in 2001
  - now different versions available in
    72+ regions/editions and 30+ languages
  - English-language versions cover the last 30 days' articles from 4,500+ sites
  - it automatically gathers, clusters and ranks articles
  - but humans select the source sites
    - accusations of bias, censorship, failure to censor, illegal reuse,...

Google News: How it Works

- Gathers the articles from the different sites
- Clusters them using agglomerative hierarchical clustering
- Ranks them on the basis of:
  - timeliness of the article
  - whether it is an original piece
  - placement by the editors on the source page
  - apparent scope and impact
  - popularity (click rate of the article by users of Google News)
  - authority of the source based on
    - number of non-duplicate stories
    - length of articles
    - breadth of articles
    - number of important/breaking stories
    - click rate of the site by users of Google News
    - average quality of the writing
    - PageRank of the source's web site
    - real world data on the news source (e.g. number of employees)

Google News: Customization

- User can choose:
  - edition (e.g. Ireland)
  - page style (e.g. Modern, Classic)
  - the importance of standard sections
  - new sections (based on keywords)

Google News: Personalization

- Recommended stories
  - different recommendations for different users
- It records your selections
  - hence, when using Google News you must sign-in to you Google account
  - the more you use Google News in this fashion, the better the recommendations become
- It recommends stories to you that similar users have selected
  - a similar user is one who has selected stories similar to the ones you have selected
  - A. Das et al (2007): GoogleNews Personalization: Scalable Online Collaborative Filtering, Procs. of the 16th International World Wide Web Conferences, pp.271-280
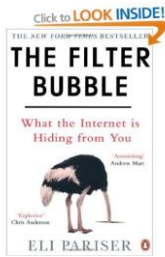
The Filter Bubble

The Filter Bubble

"a unique universe of information for each of us"

- A negative feedback loop:
  - the technology shows us only what it thinks we will like
  - we only consume items from those that it shows us
  - the technology treats this as confirmation that it is correct, and so the process repeats

A Counter-Argument

"Eli has a fundamental misunderstanding of what personalization is, leading him to the wrong conclusion. The goal of personalization and recommendations is discovery.

"Recommendations help people find things they would have difficulty finding on their own."

Adaptive systems seek to extend our tastes or increase our knowledge, not to limit them

Greg Linden