# CHAPTER 15: Information Search

*Designing the User Interface:*

*Strategies for Effective Human-Computer Interaction*

*Sixth Edition*

Ben Shneiderman, Catherine Plaisant,
Maxine S. Cohen, Steven M. Jacobs, and Niklas Elmqvist

*in collaboration with*

*Nicholas Diakopoulos*

**Addison Wesley**
is an imprint of

PEARSON

# Introduction to Information Search

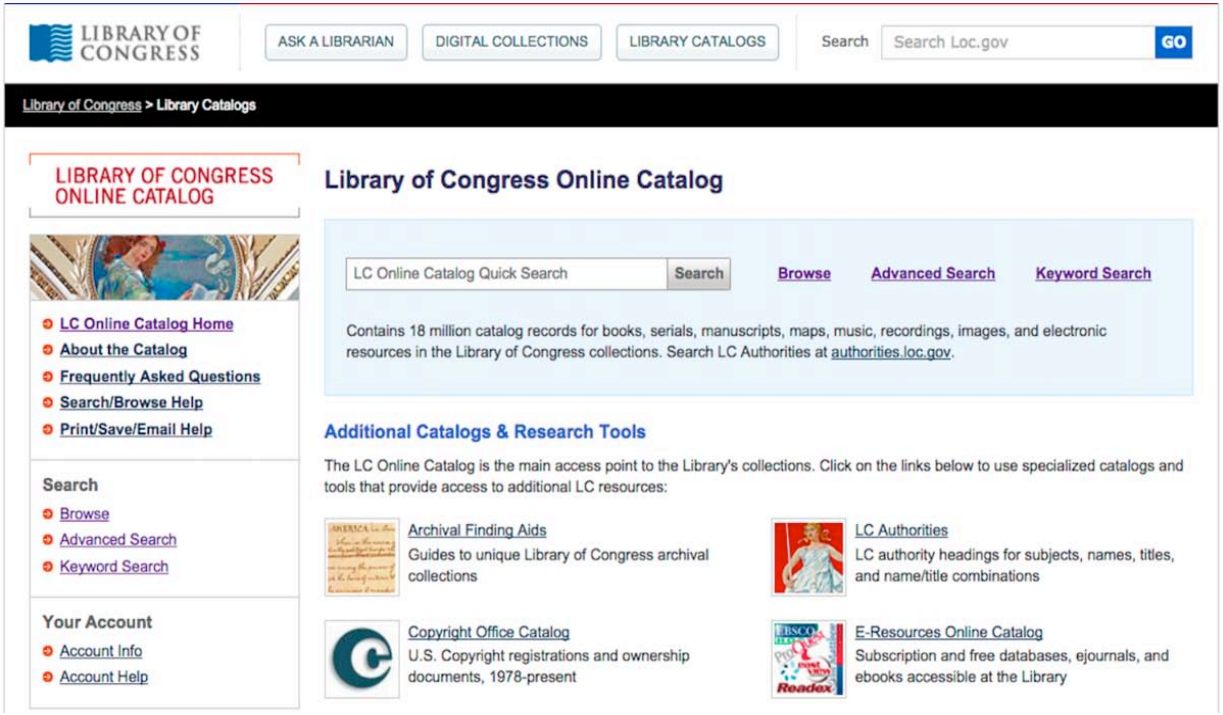
## Topics

1. Introduction
2. Five-phase search framework
3. Dynamic queries and faceted search
4. Command languages and “natural” language queries
5. Multimedia Document Search & specialized search
6. The Social aspects of search

# Introduction to Information Search (continued)

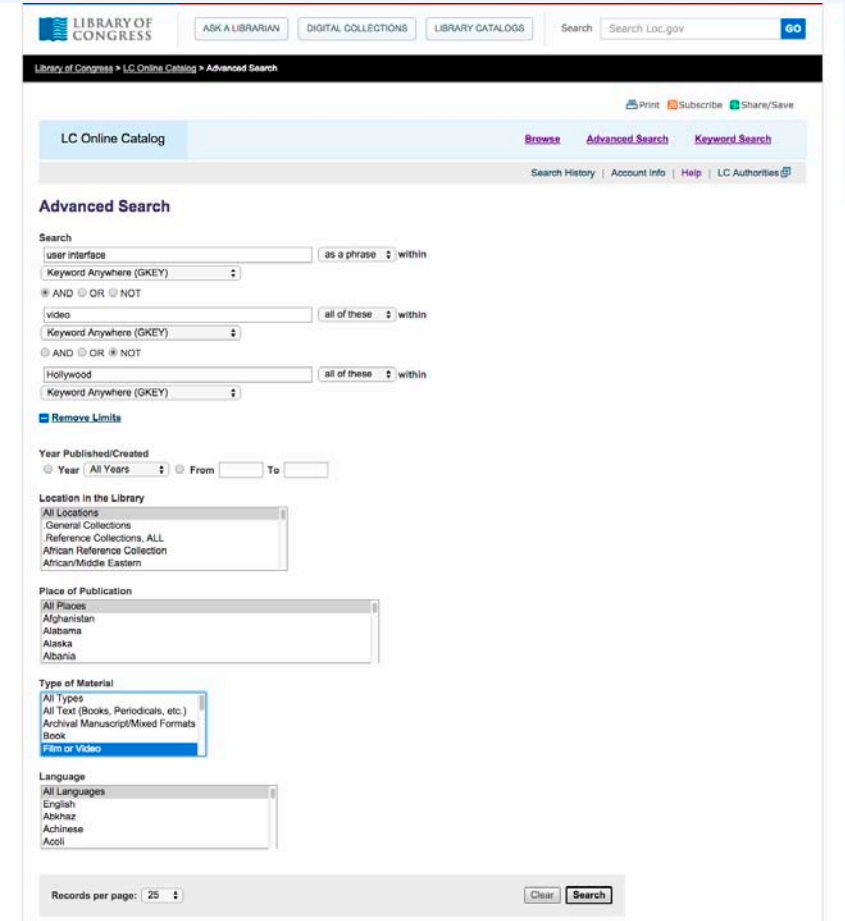
- Information search should be a joyous experience, but it still can take a skilled searcher with robust tools to perform an effective search
- Information retrieval and database management have evolved into:
  - Information seeking, filtering, collaborative filtering, sensemaking, and visual analytics.
  - Alternating these strategies is called “berry picking”
- All the above is complicated by the increased volume of material to search
  - Data mining
  - Deep learning

# Introduction to Information Search (continued)




- The home page of the U.S. Library of Congress Online Catalog (catalog.loc.gov) shows the simple search box prominently placed at the top of the page, and provides alternative means of finding items of interest in the diverse collections
    - Advanced search interfaces are provided to accommodate experienced searchers

# Introduction to Information Search (concluded)




- The advanced search interface of the U.S. Library of Congress Online Catalog (catalog.loc.gov)
  - The entire page is now dedicated to search controls and tips
  - Using checkboxes, text fields and menus users can compose Boolean queries, restrict the search scope to a subset of the collections, and apply filters based on metadata
  - Regular users sign up for an account to save results and keep a search history to facilitate re-finding

# Search terminology

- *Task objects* (such as movies for rent) are stored in structured relational databases, textual document libraries, or multimedia document libraries
- A *structured relational database* consists of *relations* and a *schema* to describe the relations
- *Relations* have *items* (usually called *tuples* or *records*), and each item has multiple *attributes* (often called *fields*), which each have *attribute values*
- A *library* consists of a set of *collections* (typically up to a few hundred collections per library) plus some *descriptive attributes* or *metadata* about the library (for example, name, location, owner)

# Search terminology (continued)

- *Digital libraries* are generally sets of carefully selected and cataloged collections
  - *Digital archives* tend to be more loosely organized
- *Directories* hold metadata about the items in a library and point users to the appropriate locations
  - for example, the NASA Global Change Master Directory helps scientists locate datasets in the many NASA's archives
- Items in *unstructured collections* like the web have no (or very few) attributes

# Search terminology (continued)

- *Task actions* are decomposed into *browsing* or *searching*
- Here are some examples of task actions:
  - *Specific fact finding (known-item search)*
    - Find the e-mail address of the President of the United States.
  - *Extended fact finding*
    - What other books are by the author of “Jurassic Park”?
  - *Exploration of availability*
    - Is there new work on voice recognition in the ACM digital library?
  - *Open-ended browsing and problem analysis*
    - Is there promising new research on fibromyalgia that might help my patient?

# Search terminology (concluded)

- Once users have clarified their information needs, the first step towards satisfying those needs is *deciding where to search*
- Supplemental *finding aids* can help users to clarify and pursue their information needs, e.g. table of contents or indexes
- Additional *preview and overview surrogates* for items and collections can be created to facilitate browsing

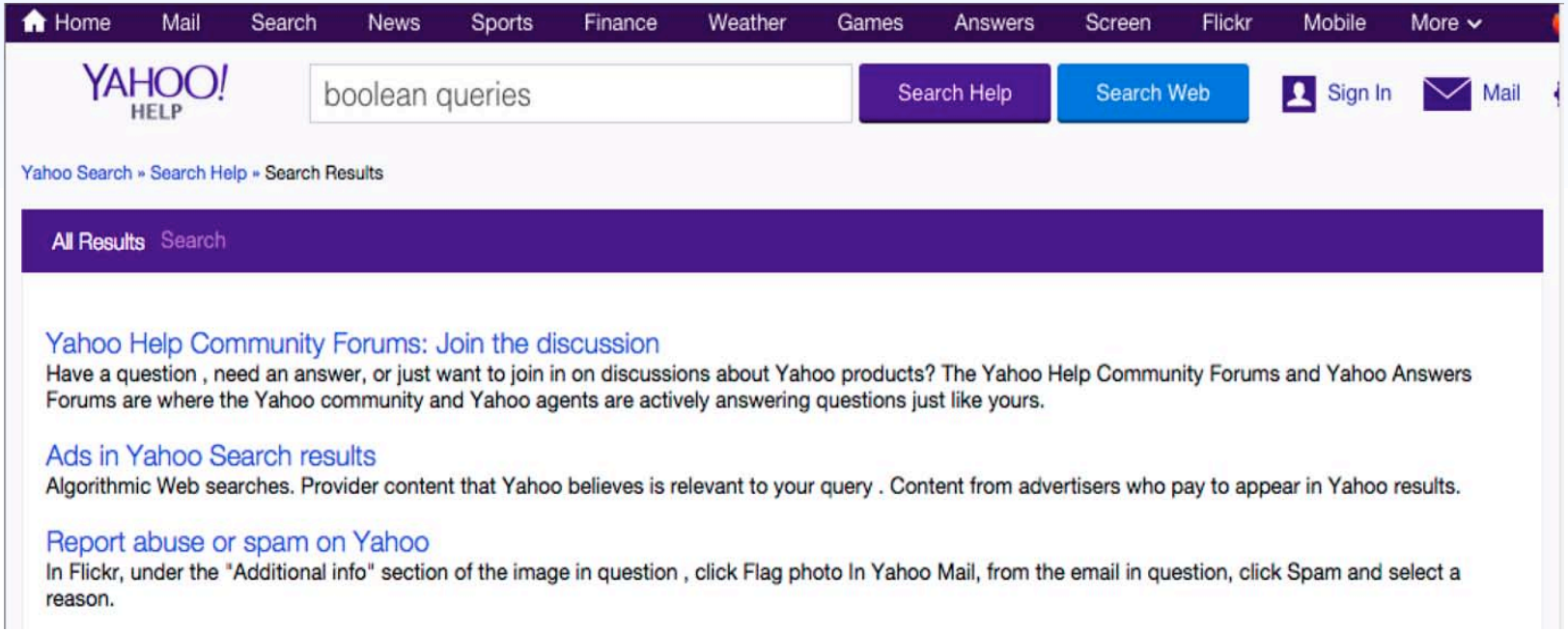
# Five-phase framework for search user interfaces

1. *Formulation:* expressing the search
2. *Initiation of action:* launching the search
3. *Review of results:* reading messages and outcomes
4. *Refinement:* formulating the next step
5. *Use:* compiling or disseminating insight

# Five-phase framework for search user interfaces (concluded)

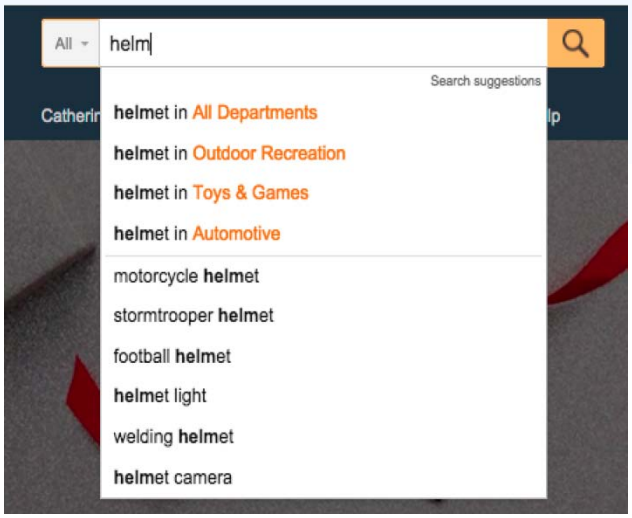
1. *Formulation*
   - Use simple and advanced search
   - Limit the search using structured fields such as year, media, or location
   - Recognize phrases to allow entry of names, such as George Washington
   - Permit variants to allow relaxation of search constraints (e.g. phonetic variations)
   - Control the size of the initial result set.
   - Use scoping of source carefully
   - Provide suggestions, hints, common sources
2. *Initiation of action*
   - Explicit actions are initiated by buttons with consistent labels (such as “Search”)
   - Implicit actions are initiated by changes to a parameter and update results immediately
   - Guide users to successful or past queries with auto-complete
3. *Review of results*
   - Keep search terms and constraints visible
   - Provide an overview of the results (e.g. total #)
   - Categorize results using metadata (by attribute value, topics, and so on).
   - Provide descriptive previews of each result item
   - Highlight search terms in results
   - Allow examination of selected items
   - Provide visualizations when appropriate (e.g. maps or timelines)
   - Allow adjustment of the size of the result set and which fields are displayed.
   - Allow change of sequencing (alphabetical, chronological, relevance ranked, etc.)
4. *Refinement*
   - Guide users in progressive refinement with meaningful messages
   - Make changing of search parameters convenient
   - Provide related searches
   - Provide suggestions for error correction (without forcing correction)
5. *Use*
   - Embed actions in results when possible
   - Allow queries, settings and results to be saved and annotated, sent to other applications
   - Explore collecting explicit feedback (ratings, reviews, like, etc.)

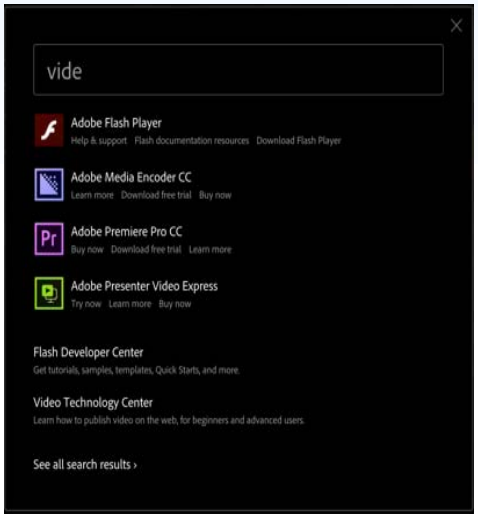
# Search formulation example




- Yahoo help search box has 2 buttons of different colors to search two different sources of information
  - Purple for searching the help information and Blue for searching the web
  - Pressing the purple button “scopes” the results to the help information only and shows results below a purple banner
  - Searching the web jumps to a different page (the normal search) that reuses the blue button color, helping users keep track of which

# More Search formulation examples

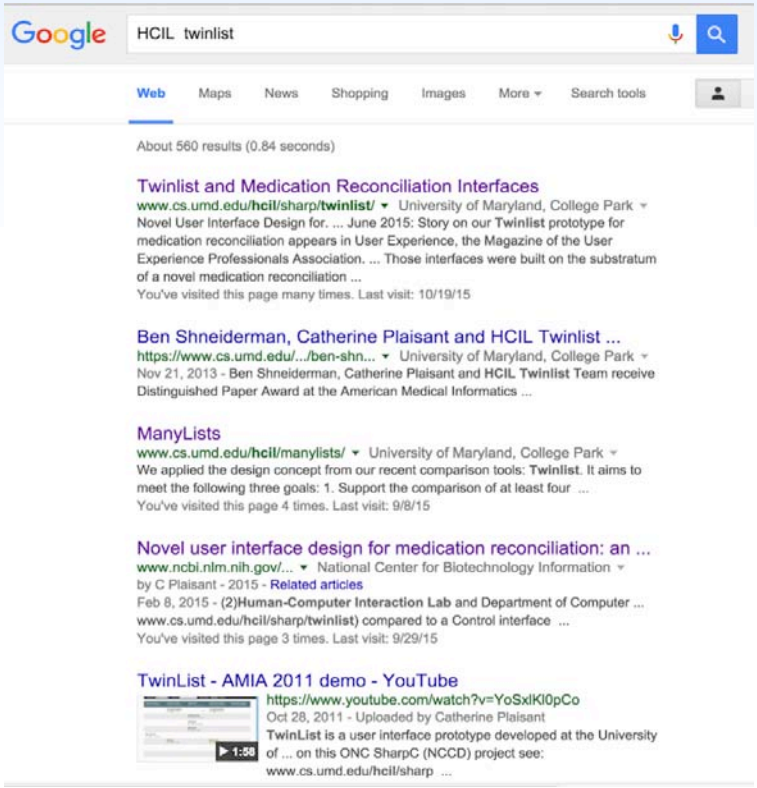




a) Apple iPhone contact manager, b) Amazon.com c) Adobe webpage

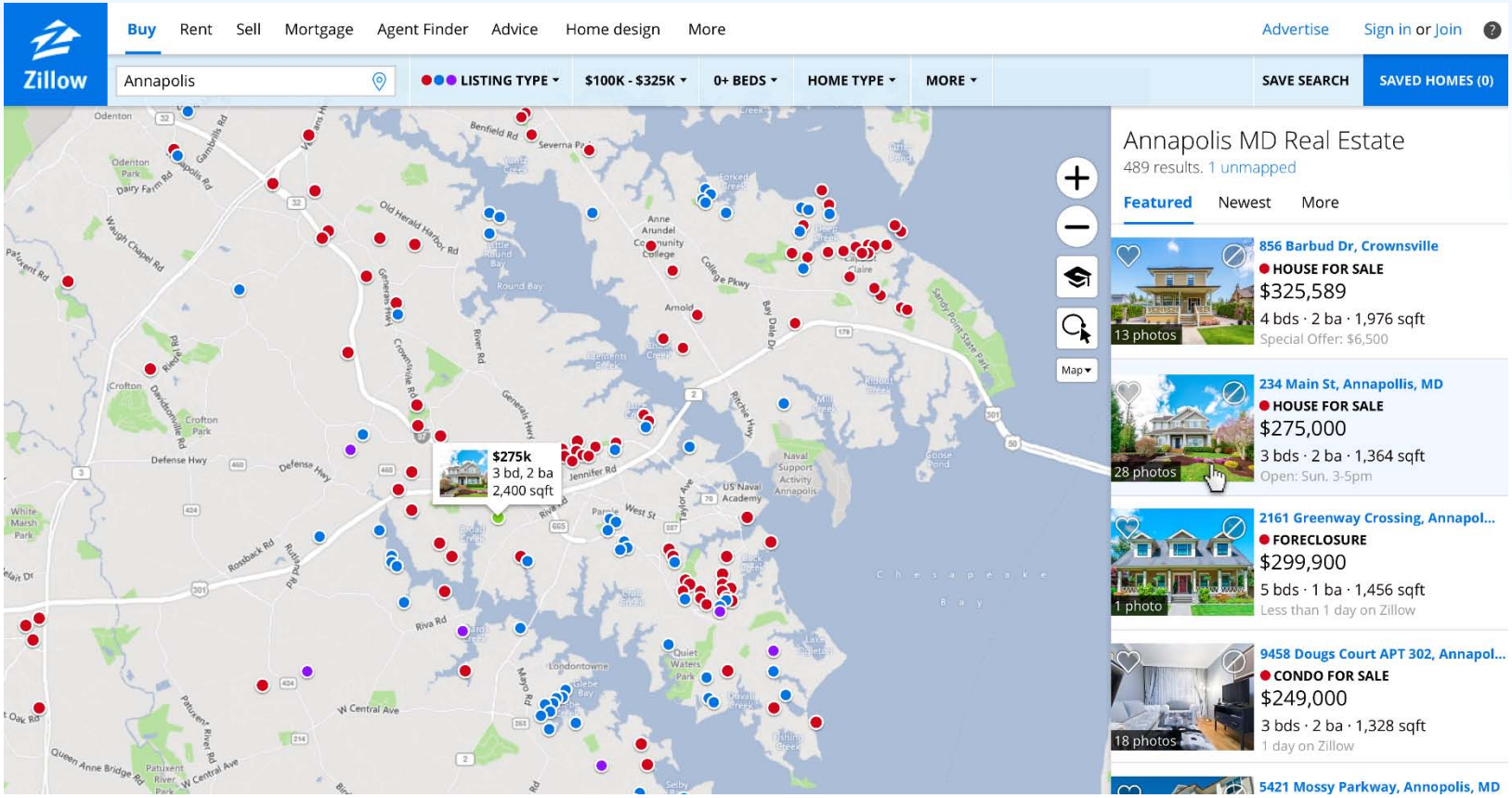
- Autocomplete suggestions can speed data entry and guide users toward successful queries
  - a) In a mobile phone address book typing one character filters the list to all names that contain that character, and the lists is updated continuously as users type
  - b) Typing “helm” in Amazon’s search box shows suggestions for “helmet light” or “welding helmet” but also suggestions to narrow the scope of the search to relevant departments
  - c) In the Adobe website suggestions include products, e.g. typing the beginning of the word “video” suggests several video editing tools

# Review of results example



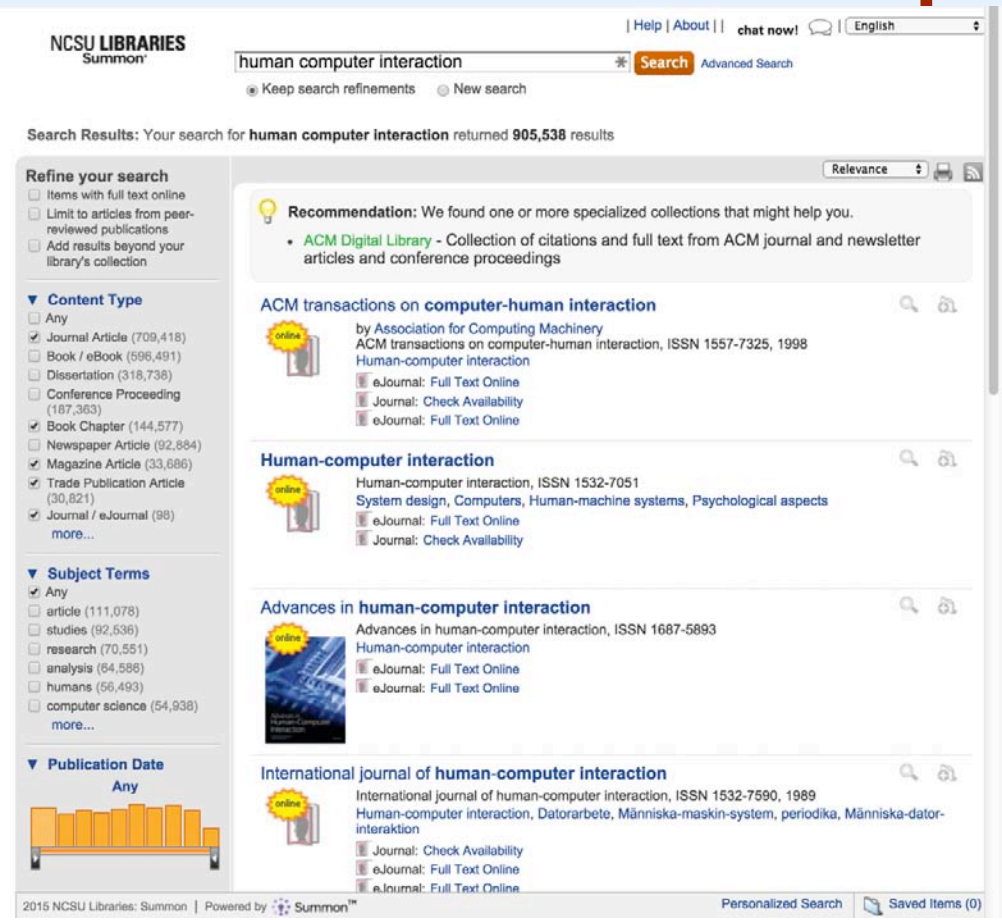

- A Google Search result list
  - A summary is provided at the top (the total number of results)
  - Each result includes preview information (or snippet)
  - Search terms are highlighted, including “Human-Computer Interaction Lab” which is the expanded variant of the search term HCIL
  - The name of the top-level organization was added (here “National Center for Biotechnology Information”) to help users judge the trustiness of the information

# Another review of results example



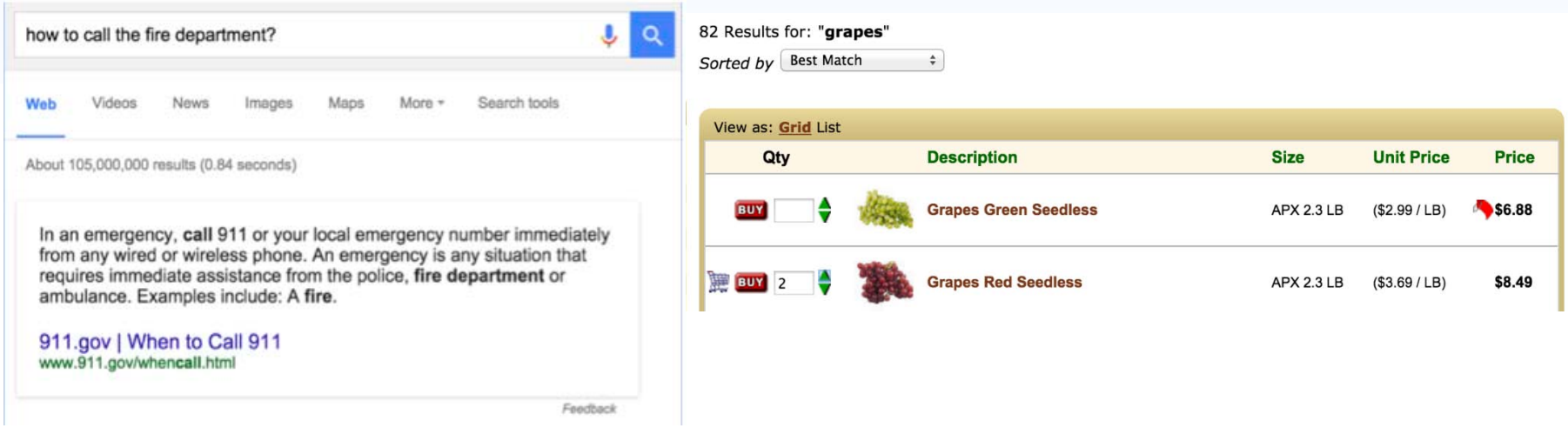

- Searching for Annapolis on the real estate website Zillow returns a list of houses and dots displayed on a map.
  - The two windows are coordinated; when the cursor hovers over a house in the result list, the location of the house is indicated on the map.
  - A click on the house would bring all the details displayed in an overlapping window.

# Another review of results example




- A search for “user interface” powered by SummonTM for a university library catalog returns a very large number of results
  - On the left users can see the number of results for categories organized by Content Type, Subject Terms, or Publication dates. It provides an overview of the results, reveals how the search was done (e.g. here the default search does not return dissertations) and facilitates further refinement of the search.
  - The menu at the upper right allows users to sort results by relevance or by date. Help is available with a “Chat now” button, to chat with a librarian.
    (http://www.lib.ncsu.edu)

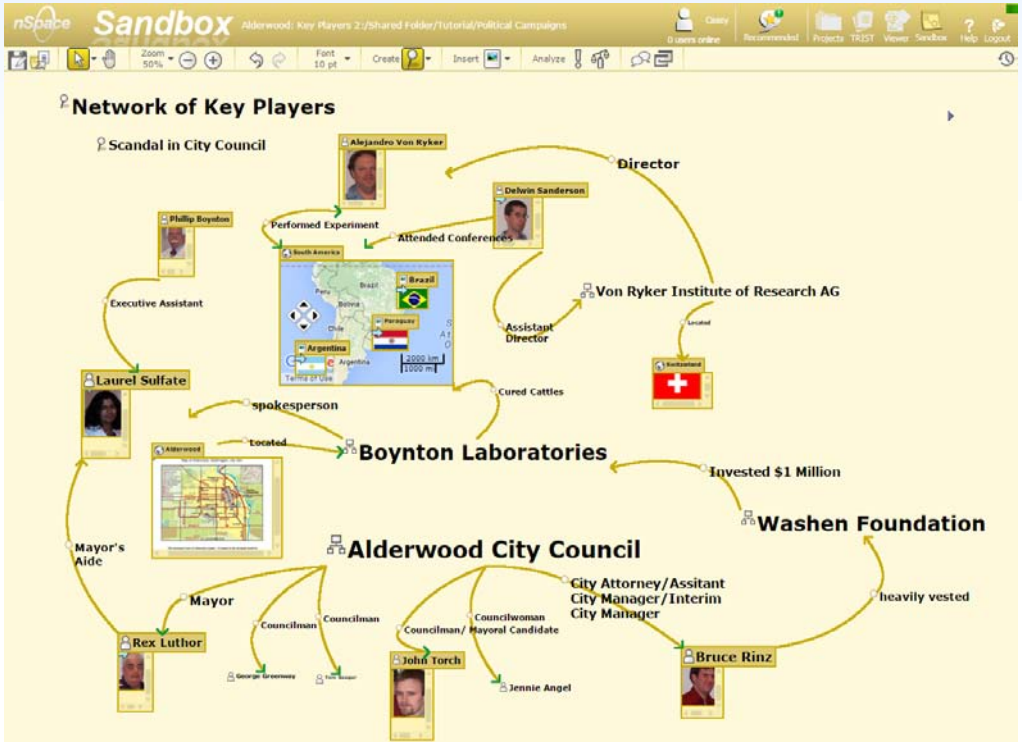
# Use of the results example




left: Google Search, and right: Peapod.com

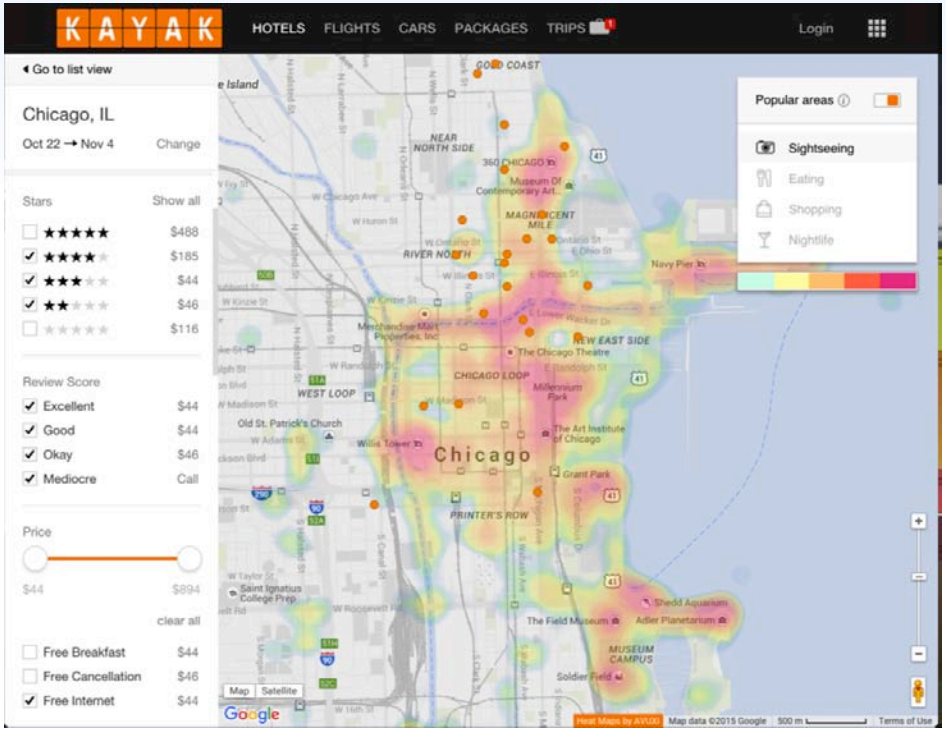
- When possible (and important) provide information or simple actions without requiring users to leave the search results page
    - On the left users get the answer to their safety critical question at the top of the result list
    - On the right shoppers looking for groceries can specify quantity and buy directly from the list of results after a search on “grapes”

# Another use of the results example



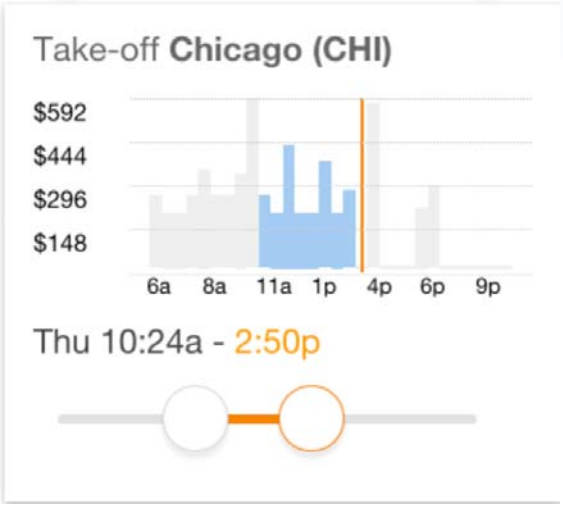

- nSpace Sandbox®, from Uncharted Software™ allows multiple analysts to organize and present the evidence gathered from research
- A variety of tools such as node and link diagramming, automatic source attribution, recursive evidence marshalling, timeline construction, etc. provide support for analysis and reporting

# Dynamic queries and faceted search for structured collections (example)



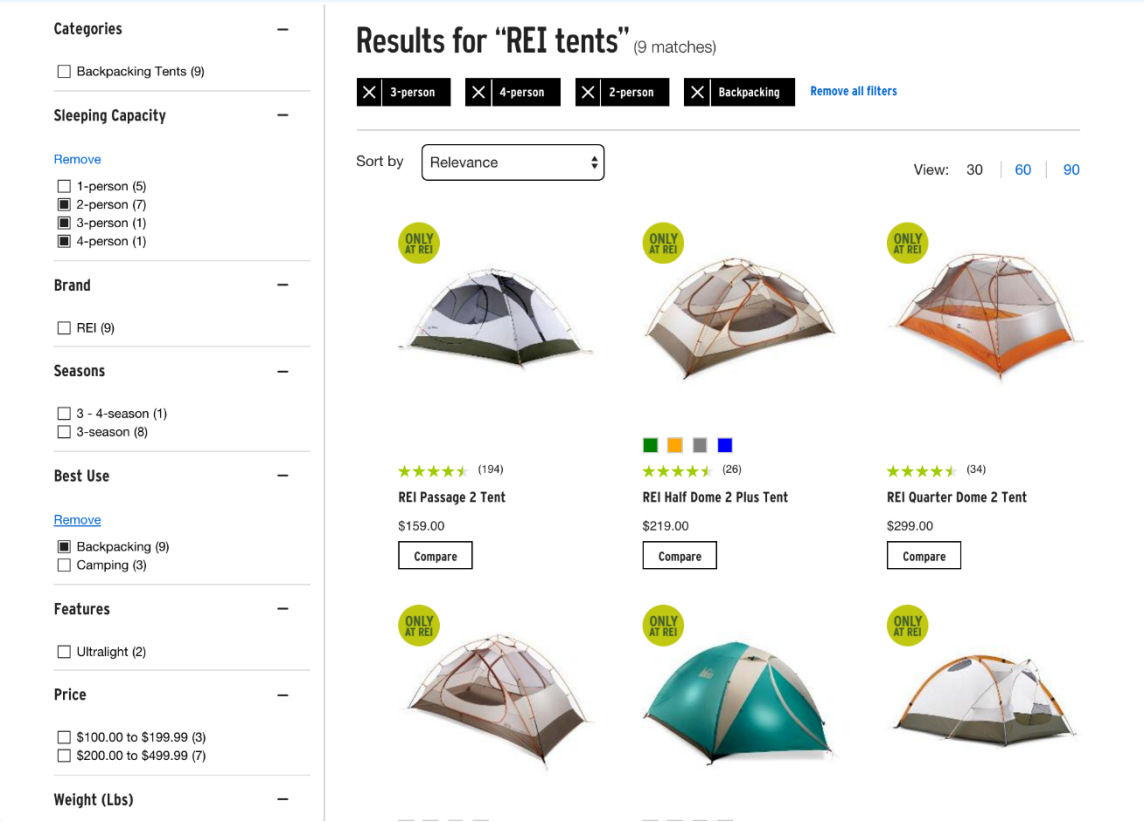

- The hotel search interface of the Kayak travel website
  - After using a form fill-in to provide the location (Chicago) and dates results are displayed in a traditional list or a map
  - The map provides an overview of the location of the hotels and can be zoomed to narrow the results. It was also augmented with a visualization of the popular sightseeing areas.
  - On the left menus are available to narrow down the categorical values, and sliders for numerical values
  - Price is important so the average price is provided for each category values

# Dynamic queries and faceted search for structured collections (example concluded)



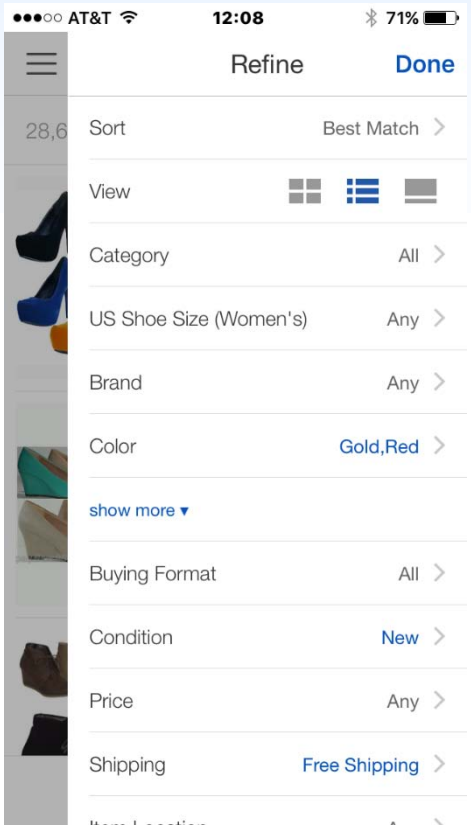

- A preview of the price of available flights guides users narrow down the time range for take-off
- The preview eliminates empty result sets, and avoids high expenses

# Faceted search example




- Faceted search interface of REI
  - Here users searched for "REI tents" and then browsed different tents by selecting values for multiple categories
  - The selected filters are clearly indicated at the top with black background, making easy for users to review the constraints and remove them

# Another faceted search example



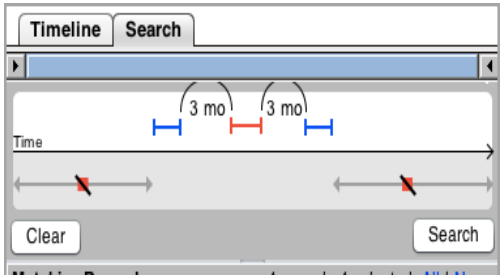

- After searching for “Golbeck” users can scroll through the results, or use the filter menu which slides to the left and partially overlaps the result list
    - “Filter (3)” indicates that three filters have already been applied (e.g. paperback as format), reducing the results to 43

# Command languages and “natural” language queries (example)

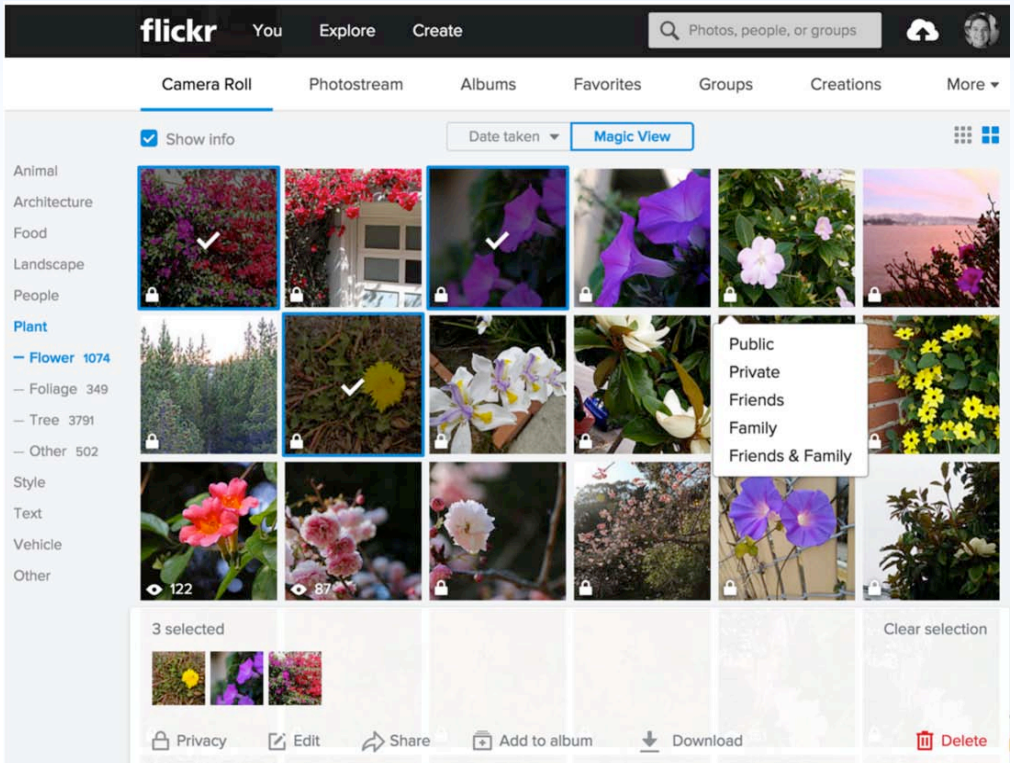
```
SELECT DOCUMENT#
FROM JOURNAL-DB
WHERE    (DATE >= 2014 AND DATE <= 2017)
  AND   (LANGUAGE = ENGLISH OR FRENCH)
  AND   (PUBLISHER = ASIST OR HFES OR ACM)
```

# Multimedia Document Search and other specialized searches

- Image search
- Video search
- Audio search
- Geographic information search
- Multilingual search
- Other specializes searches

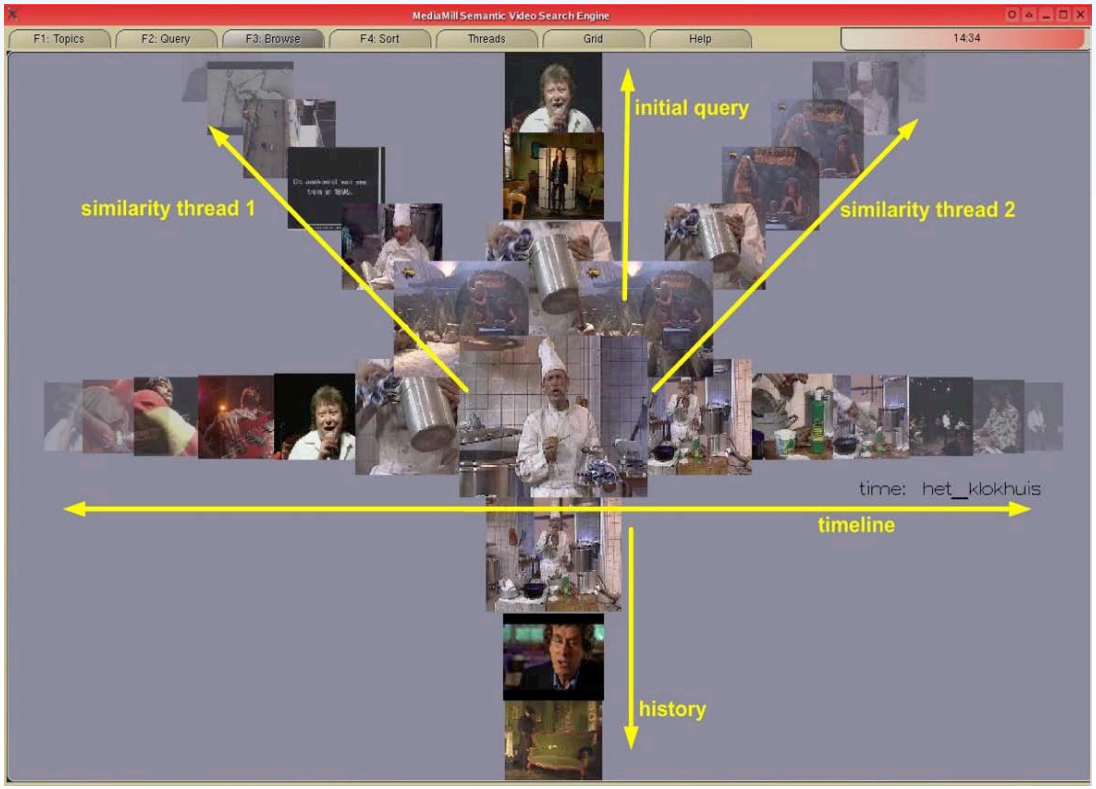


# Multimedia Document Search (example)




- The “Magic View” of Yahoo photos automatically generates topic tags for each photo
  - Here users selected the photos with flowers
  - Three photos are selected and ready to be shared
  - The privacy setting is visible and can be changed with a menu

# Another Multimedia Document Search example

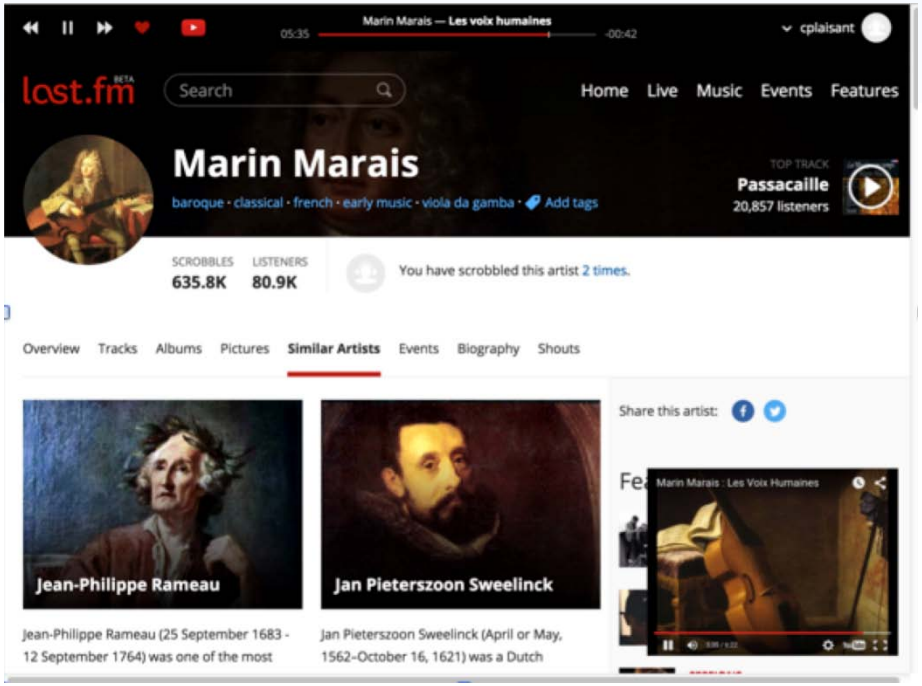



- The ForkBrowser of the MediaMill semantic video search engine (de Rooij, 2008) which allows the user to browse the video collection along various dimensions exploring different characteristics of the collection

# The social aspects of search

- *Social search* as “an umbrella term” describing search acts that make use of social interactions with others
  - May be explicit or implicit, co-located or remote, synchronous or asynchronous
  - Social bookmarking and ranking, e.g. Reddit
  - Personalized search built on user profiles, e.g. past site visits
  - Collaborative filtering and recommender systems, e.g Netflix
  - Music recommendation, e.g. Pandora

# The social aspects of search (example)




- Last.fm is an example of online radio using playlists created automatically
- The process starts by users selecting a start point (e.g. a song or artist they like) then users provide feedback on the suggestions by clicking on the heart or skipping the track