Information Architecture and Search

INFO 200

Part II

Joseph Janes
Associate Professor, Information School

Information School
UNIVERSITY of WASHINGTON

W

Information Architecture & Search 2
agenda

- search and what makes it work: unstructured search
  - information organization
  - controlled vocabulary
- and what about Google?

---

search

search is the **matching** of **representations** in a **database** by means of an **algorithm**

(representations = **metadata**)

this implies those representations, databases, algorithms have been somehow **created, defined, decided on** (all information behaviors)

and **structured** – *but what if they're not? or at least not in the same way?*

this more complex instance is referred to as unstructured search

no metadata, no search

these are all examples of **information systems** – so where is **power**?

---

information architecture/organization

let's start here: Washington State Legislature (to find bills)

Rosenfeld et al: "We organize to **understand**, to **explain**, and to **control**."

"We organize information so that people can **find** the right answers to their questions, and to give them **context** to understand those answers. context

"We strive to support casual **browsing** and directed **searching**. not the same things at all

"Our aim is to **design** organization and labeling systems that **make sense to others**." you are not them

"Our classification systems inherently reflect **social and political perspectives and objectives**." *inherently?*

what we call something can often become how we think about it

classification

Grouping things together based on their shared qualities

These groups form a taxonomy, which is often hierarchical
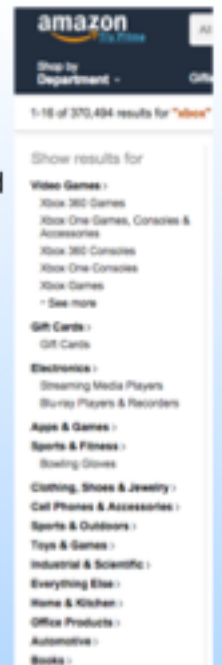
For example: Amazon

faceted classification

Organizing by multiple, parallel taxonomies

Each taxonomy considers a different quality (facet) and acts as an independent filter

Movies: genre, length, rating, directors, actors, distributor, etc. IMDB

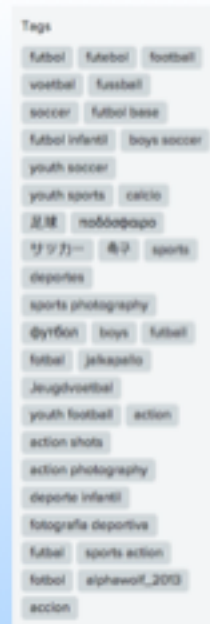**Clothing**: gender, function, fabric, color, etc.

Classes: level, credits, time, area of knowledge, size, etc.

folksonomy

Allowing end-users to tag items with various categories or descriptive words

What are the benefits and drawbacks of this approach?

controlled vocabulary

these are all examples – **constructed** (decided, designed) set of words and phrases that **help people organize, browse, and search data**

authority file/thesaurus

what do you call people who do what I do for a living?

faculty (canonical, preferred form)
- teachers
- educators
- instructors
- professors
- lecturers
- coaches
- guides
- …

synonymy

disambiguation list

works both ways

pitch
- pitch (throw)
- pitch (propose)
- pitch (assemble)
- pitch (playing field)
- pitch (frequency)
- pitch (resin)
- pitch (angle)
- pitch (distance)

polysemy

so what about Google?

what about Google indeed…first some search examples Google
where is the metadata? (**no metadata, no search**, remember)
   signals
   and they're not the only game in town - - -

Amazon Alexa and the Search for the One Perfect Answer

Voice computing seeks to deliver a single correct response to any query. That's why it's going to upend our relationship with information.

WIRED

How Voice Computing Will Transform the Way We Live, Work, and Think

Talk to Me

James Vlahos

This article is adapted from Talk to Me: How Voice Computing Will Transform the Way We Live, Work, and Think, by James Vlahos, to be published in March by Houghton Mifflin Harcourt. HOUGHTON MIFFLIN HARCOURT

According to one market survey, people ask their smart speakers to answer questions more often than they do anything else with them. Tunstall-Pedoe's vision of computers responding to our queries in a single pass—providing one-shot answers, as they are known in the search community—has gone mainstream. The internet and the multibillion-dollar business ecosystems it supports are changing irrevocably. So, too, is the creation, distribution, and control of information—the very nature of how we know what we know.

The second component of the system amassed facts. Unlike a search engine, which simply pointed users toward websites, True Knowledge aspired to supply the answers itself. It needed to know that the population of London is 8.8 million, that LeBron James is 6'8", that George Washington's last words were "'Tis well," and so on. The great majority of these facts were not manually keyed into the system, that would have been too arduous. Instead, they were automatically retrieved from sources of structured data, where information is listed in a computer-readable format.

so what about Google?

some deeper observations:

- you don't pay for it
- you are not their customer, you are their product
- opaque to human understanding or intervention
- change the way we search, change the way we think

Information Architecture and Search

INFO 200

Part II

Joseph Janes
Associate Professor, Information School

Information School
UNIVERSITY of WASHINGTON

W