Improving the search experience by organizing Web search results with meaningful and stable categories

Bill Kules and Ben Shneiderman
University of Maryland, College Park

June 1, 2006
HCIL Symposium and Open House
Outline

• Motivation and examples
• Comparative study
• Design principles
• Future work

Communications of the ACM, April 2006
<table>
<thead>
<tr>
<th>Search Result</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median</td>
<td>The median is the middle of a distribution: half the scores are above the median and this makes it less sensitive to extreme scores than the mean. [1]</td>
</tr>
<tr>
<td>Mean and Median</td>
<td>Again, this statistic is easy to determine because the median literally is... Comparing the mean to the median for a set of data can give you an idea how... [2]</td>
</tr>
<tr>
<td>Introduction to Statistics: Mean, Median, and Mode</td>
<td>The goal of this lesson is to introduce the concepts of mean, median and mode. The Mean and Median Activity lets students explore mean and median in an interactive way. [4]</td>
</tr>
<tr>
<td>Mean, Median and Mode Discussion</td>
<td>Student: When do we use mean and when do we use median? ... Unfortunately often people call mean, median and mode by the same name: average. ... [5]</td>
</tr>
<tr>
<td>CancerGuide: The Median Isn't the Message</td>
<td>The text of an essay by Gould on the subject of cancer and statistics, based on...</td>
</tr>
</tbody>
</table>
Comparative study

• How can categorized overviews change
  – Search tactics
  – Search outcomes
  – Search experience
• Scenario/task: Journalist asked to generate story ideas for an upcoming meeting
  – Topics like “human smuggling”
Experimental conditions

Top 100 results containing urban sprawl

Stopping Sprawl Main Page - Sierra Club
The Challenge to the Sprawl Campaign works to fight poorly planned runaway development and promotes smart growth communities.
www.sierraclub.org/sprawl/

Sprawl Overview - Sierra Club
Sprawl destroys more than two million acres of parks, farms and open space each year. Hundreds of urban, suburban and rural neighborhoods are using ...
www.sierraclub.org/sprawl/overview/

Urban Sprawl @ nationalgeographic.com
Loads of information about urban sprawl, including essays, books, and current news. This site examines the roots of the problem.
magma.nationalgeographic.com/ngm/data/2001/07/01/html/ft_20010701.3.html

Urban Sprawl: the Big Picture
NASA Science News: Satellites provide a "big picture" perspective on the environment. urban sprawl. But most city planners don.
science.nasa.gov/headlines/y2002/11oct_sprawl.htm

Sprawl City
Covers consumption growth and population growth and their roles in urban sprawl.
www.sprawlcity.org/

Urban sprawl - Wikipedia, the free encyclopedia
Publicly-editable encyclopedia entry supplies examples and arguments for and against urban sprawl.
en.wikipedia.org/wiki/Urban_sprawl
Experimental conditions

**Stopping Sprawl Main Page - Sierra Club**
The Challenge to the Sprawl Campaign works to fight poorly planned runaway development and promotes smart growth communities.
www.sierraclub.org/sprawl/

**Sprawl Overview - Sierra Club**
Sprawl destroys more than two million acres of parks, farms and open space each year. ... Hundreds of urban, suburban and rural neighborhoods are using ...
www.sierraclub.org/sprawl/overview/

**Urban Sprawl @ nationalgeographic.com**
Loads of information about urban sprawl, including essays, books, and current news. Looking back, this site examines the roots of the problem. ...
magna.nationalgeographic.com/rgm/data/2001/07/01/html/ft_20010701.3.html

**Urban Sprawl: the Big Picture**
NASA Science News: Satellites provide a "big picture" perspective on the environmental impacts of urban sprawl. But most city planners don.
science.nasa.gov/headlines/y2002/11oct_sprawl.htm

**Sprawl City**
Covers consumption growth and population growth and their roles in urban sprawl.
www.sprawlcity.org/

**Urban sprawl - Wikipedia, the free encyclopedia**
Publicly-editable encyclopedia entry supplies examples and arguments for and against sprawl.
en.wikipedia.org/wiki/Urban_sprawl

**Urban Sprawl**
Links to articles about urban sprawl.
www.goldcanyon.com/us/urbansprawl.html
Results

• Categorized overviews were used to narrow, filter, refine, and explore

  *I loved it. I was in love with that. I wish Google had that! With 3 clicks you have 5 pieces of information...*

  *Sort of a search within a search. That was very cool.*

• Fewer, broader queries
  – Reduced cognitive effort
Results

• Provoked ideas / questions

  For the art crimes one, I saw Science and [asked], "What does that have to do with art crimes?" …and I found out that science can help solve art crimes.

• When feeling stuck
Subjective ratings

Agreement

Organize***
Explore*
Assess***

Baseline
Categorized overview

Significance: *** p<0.001, ** p<0.05, * p< 0.10
Subjective ratings

Significance: *** p<0.001, ** p<0.05, * p< 0.10
Design principles for exploratory search interfaces

- Provide categorized overviews
  - Organize by meaningful categories
  - Expose hierarchy structure
- Tightly couple overview and list
- Support multiple kinds of categories
  - Separate into visual facets
- Support multiple kinds of visual displays
- Arrange text for scanning and skimming
### Results 1 - 21 in > Country results for song titles containing road travel

<table>
<thead>
<tr>
<th>Songs</th>
<th>Artist</th>
<th>Album</th>
</tr>
</thead>
<tbody>
<tr>
<td>As We Travel Along on the Jericho Road</td>
<td>Elvis Presley</td>
<td>The Million Dollar Quartet</td>
</tr>
<tr>
<td>Before You Travel On (Lonesome Road)</td>
<td>David Houston</td>
<td>Where Love Used to Live</td>
</tr>
<tr>
<td>A Dark Road Is a Hard Road to Travel</td>
<td>Grayson &amp; Whitter</td>
<td>Roots of Bluegrass [Country Stars]</td>
</tr>
<tr>
<td>A Dark Road Is a Hard Road to Travel</td>
<td>Grayson &amp; Whitter</td>
<td>Complete Recorded Works in Chronological Order, Vol. 2: 1928-1929</td>
</tr>
<tr>
<td>A Dark Road Is a Hard Road to Travel</td>
<td>Grayson &amp; Whitter</td>
<td>The Recordings of Grayson &amp; Whitter: Recorded 1928-1930</td>
</tr>
<tr>
<td>A Hard Road to Travel</td>
<td>Union Springs</td>
<td>Next Train Headed South</td>
</tr>
<tr>
<td>Jordan Am a Hard Road to Travel</td>
<td>Norman Blake</td>
<td>The Norman &amp; Nancy Blake Compact Disc</td>
</tr>
<tr>
<td>Jordan Am a Hard Road to Travel</td>
<td>Jimmie Driftwood</td>
<td>The Westward Movement</td>
</tr>
<tr>
<td>Jordan Am a Hard Road to Travel</td>
<td>Jimmie Driftwood</td>
<td>Americana</td>
</tr>
<tr>
<td>Jordan Am a Hard Road to Travel</td>
<td>J.E. Mainer</td>
<td>20 Old-Time Favorites</td>
</tr>
</tbody>
</table>
Thank you!

For more information:

wmk@cs.umd.edu
www.cs.umd.edu/hcil/categorizedoverview

This research was partially supported by an AOL Fellowship in Human-Computer Interaction and National Science Foundation Digital Government Initiative grant (EIA 0129978) “Towards a Statistical Knowledge Network.”