User Interfaces in the Digital Library

KAN Min-Yen
HCI in the digital library

• For digital libraries, it is the means of expediting the information seeking process for a human user

“When an interactive system is well-designed, the interface almost disappears, enabling users to concentrate on their work, exploration and pleasure” – Shneiderman 97
Tenets of HCI

• Offer informative feedback

• Reduce working memory load

• Provide alternative interfaces for novice and expert users
ISP Overview

1. Information Need
2. Query
3. Send to System
4. Receive Results
5. Evaluate Results
6. Reformulate

- Done?
- Stop
Evaluating HCI

Subjective
• Ergonomics
• Subjective satisfaction

Objective
• Speed of performance
• Rate of errors
• Time to learn
• Retention over time

Dimensions
• Target Audience
  – Novice vs. Expert
  – Capabilities
• Target Task
• Context / Environment
• In situ vs. Lab environment
• Intrinsic versus Extrinsic
Outline

• Searching
  – Query formulation
  – Displaying results

• Browsing
  – Categories and Hierarchies

• Integrating Frameworks
Query specification

- **Old:** originated from command line interfaces
  - Suit the system, not the user
  - e.g. “FIND TW Mt St. Helens AND DATE 1981”

- **Then:** translated for users on OPACs
  - Subject: 

- **Now:** graphical means for query specification
VQuery (Jones 98)

Search for any documents in "Oxford Text Archive" containing either lion and tiger; or jungle; but not Tarzan
Filter / Flow model

• Users can select from the set of attributes and get an appropriate filter widget
  – (type-in for interest areas, sliders for cost, and buttons for scholarships)

• The widget is placed on the screen with flow lines showing ANDs (sequential flow) and ORs (parallel flows).

• Water flow dynamically indicates relevant # of items
Against a controlled vocabulary

• Lists
  – Seen on last slide from Filter / Flow
  – Only good if limited number of entries

• Partial Fill-in
  – Show possible completions of query terms if under a certain number (~5)

• Re-writing of form fill-in queries
  – Who is the leader of Sudan? →
    Who is the head of state of X (Sudan)?
Tiles (Anick et al. 90)

• Use table to represent conjunction and disjunction
  – Conjunction (AND): rows
  – Disjunction: columns

• Use activation to see query preview from index

<table>
<thead>
<tr>
<th>Copy</th>
<th>Backup</th>
<th>Saveset</th>
<th>From</th>
<th>Tape</th>
<th>Under</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Scratch tape</td>
<td>Version</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>5.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Version 5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>840</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Apply Changes  Display Titles
Query by Example

Query:

- QBIC System (Flickner et al. 1995)
  
  http://www.hermitagemuseum.org/cgi-bin/db2www/qbicLayout.mac/qbic?selLang=English

1) Palace in Athens
   Klenze, Leo von
   1835

2) Allegory of Rome
   Stefano, Francesco di (Il Pesellino)
   Circa 1448

3) Allegory of Carthage
   Stefano, Francesco di (Il Pesellino)
   Circa 1448

4) Boats at Saintes-Maries
   Gogh, Vincent van
   1888

5) A Game of Bowls
   Matisse, Henri
   1908

6) Lake
   Hodler, Ferdinand
   1910-1912

7) The Top of Etna from the East, View from San Vincenzo

8) View of Lake Maggiore

- What’s the corollary for text queries?
Faceted Metadata

Faceted objects give rise to easy methods for summarizing the data

- The Flamenco Project
  http://flamenco.sims.berkeley.edu/cgi-bin/flamenco/arts/Flamenco?username=default

- Facet Map
  http://facetmap.com
## Flamenco Image Search

### Media
- **Aquatint** (2025)
- **Basketry** (44)
- **Book** (666)
- **Ceramic** (1008)
- **Costume** (560)
- **Decorative Box** (163)
- **Domestic Object** (176)
- **Drawing** (2624)
- **Drypoint** (2143)
- **Etching** (3607)
- **Furnishing** (127)
- **Glass** (661)
- **More...**

### Nature
- **Animal Material** (515)
- **Birds** (1437)
- **Bodies of Water** (3604)
- **Creatures** (801)
- **Fish** (219)
- **Flowers** (1220)
- **Geological Formations** (2122)
- **Heavens** (2353)
- **Hoofed Mammals** (2480)
- **Invertebrates and Arthropods** (300)
- **Mammals** (2116)
- **Plant Material** (788)
- **More...**

### Location
- **Africa** (463)
- **Asia** (1325)
- **Australia** (21)
- **Central America** (134)
- **Europe** (23331)
- **Middle East** (78)
- **North America** (11111)
- **Oceania** (111)
- **Roman Empire** (4)
- **South America** (463)

### Places and Spaces
- **Bridges** (592)
- **Building Parts** (3088)
- **Buildings** (2393)
- **Dwellings** (1709)
- **Lawn** (20)
- **Open Spaces** (1732)
- **Roads** (1480)
- **Workplaces** (753)

### Date
- **1 - 1000 A.D.** (138)
- **12th century** (3)
- **13th century** (4)
- **14th century** (3)
- **15th century** (76)
- **16th century** (1225)
- **17th century** (3059)
- **18th century** (2287)
- **19th century** (7852)
- **20th - 21st century** (18)
- **20th century** (14298)
- **21st century** (12)
- **More...**

### People
- **Aristocrats** (974)
- **Children** (2501)
- **Men** (7372)
- **Occupations** (715)
- **Women** (3906)

### Shapes, Colors, and Materials
- **Colors** (5881)
- **Decorations** (1441)
- **Fabrics** (345)
- **Metal** (273)
- **Paper** (457)
- **Shapes** (2752)
- **Visual Framing** (5911)
Geographic queries

- Geospatial data makes the 2 & 3-D visualization a good metaphor
Geospatial queries

1. Point-in-polygon

2. Region

3. Distance and Buffer Zone

4. Path

5. Multimedia (QBE)
Keyword in Context (KWIC)

Differentiate Keyword in Context versus Query Word in Context

• QWIC: query based
• KWIC: any keyword (can be from metadata)
**Tilebars (Hearst 95)**

- Each row represents a topic (conjunction)

- The darkness of each tile in a row represent the frequency of occurrence of an item

---

*Tilebars* (Hearst 95)

- Topic 1: "drought"
- Topic 2: "ration"

was submitted to the UC Berkeley Digital Library Database where 187 matching records were found, as shown in the tilebars display below:

- State Drought Water Bank report
- California Water Plan Update, Volume 1
- California Water Plan Update, Volume 2
- California’s Continuing Drought 1987-91 report
- Dublin San Ramon Service District Urban Water Management plan
- New Updated Edition Urban Drought Guidebook report
- Regulatory Impact Assessment of The Final Water Quality report
- Urban Water Management Plan plan
- Irvine Ranch Water District 1995 Urban Water Management plan
- City of Santa Cruz Urban Water Management Plan plan
- California Water Plan Update: Executive Summary plan
- Bay-Delta Oversight Council Draft: Briefing Paper on State
Infocrystal (Spoerri 95)

- Explosion of the Venn Diagram

- Uses shape and color to model to organize results
More on Infocrystal

How to create **Nested Queries?**

- Queries can build on others to form nested queries
- Query terms can be weighted between 1 (default) and -1 (logical not)
- The icons are placed closer to the center of the crystal as they are more relevant
Cartographic Representations

- X and Y dimensions need to connote some real value for users
- Arbitrary use of X,Y plane usually not a good idea, especially if values aren’t replicable.
Scatter / Gather

Gradual refinement by selection of clusters:

**Gather**: select cluster(s) (relevance feedback)

**Scatter**: throw them into bins using automatic clustering
Scatter / Gather

Second cluster: stars (of Broadway)  Third cluster: stars (as heavenly objects)
The results display: “The big with the small”

– Overview + Details
– Focus + Context
– Magic Lenses
– Zoom / Pan
– Cue Techniques

– Brushing and Linking
– Searching with Browsing
Overview + Details
Focus + Context : “Fisheye Distortion”
Superbook

- Searching a software technical manual
- Use a expandable table of contents to show results
- Shows hits within each level as query preview
- Closely modeled by browser in Windows Explorer

<table>
<thead>
<tr>
<th>12 An Interactive Environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 Preface</td>
</tr>
<tr>
<td>1 How to Beat the Lottery</td>
</tr>
<tr>
<td>1 Tutorial Introduction to</td>
</tr>
<tr>
<td>7 Graphical Methods in S</td>
</tr>
<tr>
<td>2 Loading in Data</td>
</tr>
<tr>
<td>...</td>
</tr>
<tr>
<td>3 Building Plots</td>
</tr>
<tr>
<td>2 Specialized Plots</td>
</tr>
<tr>
<td>2 Advanced Use of S</td>
</tr>
<tr>
<td>2 Appendix I</td>
</tr>
</tbody>
</table>
Summarization atop IR

• Like Superbook, but trades focus for extractive summary (Kan et al. 02)
  – **Relevant**: information for summary
  – **Irrelevant**: too broad or not on topic, for broader queries
  – **Intricate**: too detailed, for follow-ups

• Furnas’ “Degree of Interest”
Focus + Context in NetBeans IDE

```java
protected void show_scoreSheet (Player p, int gameNum) {
...
}

protected void enable_buttons (Player p, int gameNum) {
    for (int but = CellCodes.ONES; but <= CellCodes.CHANCE; but++) {
        if (p.getScoreCell(gameNum, but) < 0)
            cats[but].setEnabled(true);
    }
}

protected void disable_AllCells () {
...
}

protected void reset_board() {
...
}

protected void clickScoreCell(int code, Match m) {
...
}

private void attach_listener (final Button b, final int code, final Match m) {
...
}

private TextField[] make_five_fields (int row) {
...
}

private void oneRow (final String label, final int butCode, final Match m) {
...
}

private void make_totals_rows(String s, int row) {
...
}
```
Tablelens (InXight)

- Focus + context to give fisheye distortion to table rows and values
Visualizing Hierarchies

- Fisheye zoom (Furnas 84)
- One implementation: Inxight’s StarTree
Multiviews via Magic Lenses

• More than two views of the same space
• Lenses show and hide information that isn’t pertinent
• Has been applied in query focusing as well
• Demo: [http://www.memorialhall.mass.edu/activities/media.jsp?itemid=15414&img=0](http://www.memorialhall.mass.edu/activities/media.jsp?itemid=15414&img=0)
Contextual Cues

Depth

Window Decoration
Multiple Hierarchies: Faceted Queries

Refine your search further within these categories:

**Media (group results)**
- costume (4), domestic object (1),
- drawing (186), etching (4), metalwork (4),
- print (1)

**Location (group results)**
- all > Asia > India
- Assam, Nagaland (1), Company School (5),
- Deccan (Maharashtra School) (1), Deccan School (1),
- Deccan-hyderabad (1),
- Gujarat (3), Jain School (1), Kashmir School (1),
- Kashmir School (Afghan) (1),
- Kashmir School? (1), more...

**Date (group results)**
- 1 - 1000 A.D. (1), 13th century (4), 15th century (1), 16th century (2), 17th century (12), 18th century (49), 19th century (53), 20th century (3), date ranges spanning multiple centuries (9), date unknown (17)

**Themes (group results)**
- military (30), mortality (2), music, writing, and sport (47), nautical (9), religion (42)

**Objects (group results)**
- clothing (15), containers (4), food (6),
- fuel (1), furnishings (17), jewelry and riches (3), lightning (1), timepieces (1),
- vehicles (3), wares (1)

**Nature (group results)**
- animal material (6), birds (18), bodies of water (14), creatures (6), fish (2), flowers (9),
- geological formations (17), heavens (3),
Additional notes on faceted interface

- Multiple interfaces possible
  - Columnar style
  - Drop down list style
  - Indented ancestors

Features
- Gives partial results
- Gives multiple filters at once
- A selection drops non-matching items from results
- Allows users to define most useful criteria
- In UI, only show fields that address audience
Conclusions

• Systems show more advanced text, graphics, and animation but some are not evaluated
• Tighter integration in query and display
• Simultaneous browsing and searching support at all levels
• Display attributes dictate useful alternatives
  – To think about: mobile applications, newer UI technologies
To think about

• What types of information do these various UI need to calculate at run-time and what types can be pre-computed?

• How do these UI support the tenets we mentioned at the beginning?

• Do you feel that HTML / WWW has enhanced UI design or deterred its creativity?

• For images and audio:
  – When is query by example a useful technique and when is it inferior to a search using metadata?