Human versus Machine in the Topic Distillation Task

Mingfang Wu
Alistair McLean
Ross Wilkinson

Gheorghe Muresan
Muh-Chyun Tang
Yuelin Li, Hyuk-Jin Lee
Nicholas J. Belkin
Outline

- Introduction to the topic distillation task:
  - Automatic vs. interactive
- Our research questions:
  - Human vs. machine performance
  - Does machine help user?
  - Linear vs. hierarchical layout
- Experimental design
- Experimental results
- Summary
Topic Distillation Task

- Topic distillation:
  - Given a broad query, finding a list of key resources.

- To be a “key resource”, a page should be a good entry point to a website which:
  - Provides credible information on the topic;
  - Is principally devoted to the topic (not too broad); and
  - Is not part of a larger site also principally devoted to the topic (not too narrow).
Task Example: *Find government information on adoption procedures*

[http://www.courtinfo.ca.gov/selfhelp/family](http://www.courtinfo.ca.gov/selfhelp/family) too broad
Task Example: *Find government information on adoption procedures*

http://www.courtinfo.ca.gov/selfhelp/family/adoption: just right
Task Example: Find government information on adoption procedures

http://www.courtinfo.ca.gov/selfhelp/family/adoption/stepparent.htm: too narrow
Task Example: *Find government information on adoption procedures*

http://www.courtinfo.ca.gov/

http://www.courtinfo.ca.gov/selfhelp/

http://www.courtinfo.ca.gov/selfhelp/family/

http://www.courtinfo.ca.gov/selfhelp/lowcost/

http://www.courtinfo.ca.gov/selfhelp/family/adoption/

(This website should meet the three conditions.)

http://www.courtinfo.ca.gov/selfhelp/family/adoption/stepparent.htm
Task:
<top>
<num>
Number: TD
<title>
adoption procedures
</title>
<desc>Description:
Information regarding how one adopts children?
</top>

Answer:

Administration for Children and Families (USHHS)
www.acf.dhhs.gov
www.acf.dhhs.gov/programs/cb/dis/afcars/
cpru.usda.gov/

California Courts Self-Help Center
www.courtinfo.ca.gov/selfhelp/family/adoption/

Mesa County Dept. of Human Services
www.mcdss.co.gov

Texas Department of Health, Central Adoption
www.tdh.texas.gov/bvs/car/car.htm
Task:

Cotton Pathology Research Unit
cpru.usda.gov/

FAS Cotton Group
ffas.usda.gov.cots/cotton.html

U.S. Cotton Data Sets
wizard.arsusda.gov/cotton/ars2.html

USDA Cotton Program
www.ams.usda.gov/cotton/

Safety and Health Topics: Textiles
www.osha-slc.gov/SLTC/textiles/
Why an interactive task?

- Information tasks are a partnership between people and their information systems;
- To understand task effectiveness, it is necessary to study not only system effectiveness, but also the effectiveness of the partnership.
- Which parts are best done by people? (e.g. text understanding)
- Which parts are best done by system? (e.g. matching a query)
- What are the trade-offs in time and efficiency?
- How to make the partnership work best?
Interactive topic distillation task

- Example topic:
  **Topic:** cotton industry
  **Search task:** You are to construct a resource list for high school students who are interested in cotton industry, from growing, harvesting cotton and turning it into cloth. Your final resource list should cover all the major aspects of this topic.

- Task analysis: To assess whether a web page is a key resource page, a searcher needs to make the following judgments about the page:
  - Is it relevant? (topical relevance)
  - Is it too broad or too narrow? (depth)
  - Does it cover a new aspect? (coverage)
How to measure “performance”?

Four criteria for the quality of a resource list:

- **Accuracy**
  - **Relevance**: The page is relevant to the topic
    - 1 = Agree strongly, 3 = Neutral, 5 = Disagree strongly
  - **Depth**: Is the page too broad, too narrow, or at the right level of detail for the topic?
    - 1 = Too broad, 3 = Right level, 5 = Too narrow

- **Coverage**
  - **Coverage**: The set of saved entry points covers all the different aspects of the topic
    - 1 = Agree strongly, 3 = Neutral, 5 = Disagree strongly
  - **Repetition**: How much repetition/overlap is there with the set of saved entry points?
    - 1 = None, 3 = Some, 5 = Way too much
How to measure performance?

- More objective measures:
  - No. of query iterations
  - No. of document surrogates seen
  - No. of documents read
  - No. of documents saved
  - Actual time used
**How to measure performance:**

- **Subjective measures**
  - searchers’ *satisfaction* with the interaction
  - searchers’ *self-perception* of their task completeness
  - searchers’ *preference* of an search system/interface
Research Questions

- **RQ1**: Could a human searcher outperform a topic distillation engine?
- **RQ2**: Would a human user perform better with a topic distillation engine than with a generic content-oriented search engine?
- **RQ3**: Could a user’s performance be improved by organizing the search results in a form appropriate for the topic distillation task?
Experimental Design

Rutgers
Best match engine

CSIRO
Topic distillation engine
Experimental design

Topics are divided into blocks

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td>1234</td>
<td>4321</td>
<td>3142</td>
<td>2413</td>
</tr>
<tr>
<td>B2</td>
<td>5678</td>
<td>8765</td>
<td>7586</td>
<td>6857</td>
</tr>
</tbody>
</table>

The system and topic block combination for the first four searchers

<table>
<thead>
<tr>
<th></th>
<th>System I: B1a</th>
<th>System II: B2a</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S2</td>
<td>System I: B2a</td>
<td>System II: B1a</td>
</tr>
<tr>
<td>S3</td>
<td>System II: B1a</td>
<td>System I: B2a</td>
</tr>
<tr>
<td>S4</td>
<td>System II: B2a</td>
<td>System I: B1a</td>
</tr>
</tbody>
</table>
Experimental Procedure

- Tutorial and demo
- For each system:
  - Hands-on practice
  - Search on the topic
- For each topic:
- Entry questionnaire
- Pre-search questionnaire
- After-search questionnaire
- After-system questionnaire
- Exit questionnaire
Research Question I – Human vs. Machine

Hypothesis

- There would be no difference in the quality of resource pages between a manually constructed and an automatically generated list.
Interfaces from Rutgers – Ranked List

Current task: 1
Assume that you are a high school student and working on an arctic exploration project. You are asked to collect some resources from the web for your project team on what kinds of exploration of the arctic are underway, especially of glaciers and ice.

Please describe your information problem:
arctic exploration

Search results:

http://energy.usgs.gov/state/ak.html
USGS Energy Program-Activities in Alaska
... that created the adjacent western Arctic Ocean Basin to the north and the Brooks Range mountain belt to the south. A full understanding of these petroleum bearing deposits and of the petroleum potential of the extensive continental shelf and adjacent ... adjacent coastal plain which borders arctic Alaska awaits an adequate exploration of the geology and tectonic history of the poorly explored western Arctic Ocean Basin and the Brooks Range. Accordingly the USGS has undertaken a program of geophysical studies and ... bedrock coring in the western Arctic Ocean Basin and a program of geologic mapping in the northeastern Brooks Range adjacent to the coastal plain of the Arctic National Wildlife Refuge. These results are of interest to U.S. exploration companies.

http://aro.aran.nova.gov/ARO_3_projects_opp.html
Arctic Research Office Projects and Opportunities
... for projects funded by the Arctic Research Initiative may be found on the website of Cooperative Institute for Arctic Research. In FY2001 NOAA received funding to provide scientific support for management decisions regarding fisheries and marine mammal interactions in the ... NOAA's Office of Ocean Exploration solicited proposals in support of its mission to expand knowledge of the ocean's physical chemical and biological environments, processes, characteristics and resources by means of interdisciplinary expeditions to unknown or poorly known oceanic regions ... to the Office of Ocean Exploration no later than 1 p.m. EST on November 7, 2001. In connection with NOAA's Ocean Exploration Initiative, NOAA held workshops July 25-26, 2001 to discuss activities that could be undertaken.

http://www.arctic.nova.us/db/register.html
Arctic theme page - Scientific
... About the Arctic Theme. http www.arctic nora webmaster arctic nora Disclaimer ...
Interfaces from CSIRO - Ranked List

   19 May 24 1999 International Adoption Availability Of Children For Nicaraguan Adoption Residents AUTHENTICATION OF ADOPTION Agencies And Time Travel Of The Nicaraguan Embassy And Consulate In U.S. Schedules Appointment With US Consulate What Documents To Bring With You... (Text starts with a list of points, possibly about adoption procedures)

2. [G33-62-2773087]
   ...the order of any Administrative Procedure Act (APA) appeal, see Editor Note The following Act is the finalized version of the 7th edition, which means the process employed by an APA appeal may have been affected by the procedures for Title 5 of the Administrative Procedure Act (APA) applicable to federal agencies. For a specific Act, procedures may be found in the Administrative Procedure Act (APA) and may be used as the administrative process involved does not render the same rule within the same time period of the Act if amended by Act 2013...

3. [G05-11-2757115]
   ...FEDERAL RULES OF CIVIL PROCEDURE WITH FORMS DECEMBER 1 2000 E PLURIBUS UNUM 3 Printed for the use of the COMMITTEE ON THE JUDICIARY HOUSE OF REPRESENTATIVES FEDERAL RULES OF CIVIL PROCEDURE December 1 2000 10TH CONGRESS COMMITTEE PRINT No 12. Session FEDERAL RULES OF CIVIL PROCEDURE WITH FORMS DECEMBER 1 2000 E PLURIBUS UNUM 3 Printed for the use of the COMMITTEE ON THE JUDICIARY HOUSE OF REPRESENTATIVES FEDERAL RULES OF CIVIL PROCEDURE December 1 2000 10TH CONGRESS COMMITTEE PRINT No 12. Session...

4. [G21-11-2287170] Administrative Procedure Act - Table of Contents
   ...Administrative Procedure Act Government Code Title 2 Division 3 Part I includes changes through 1999 legislative session. Table of Contents CHAPTER 1 Office of Administrative Hearings ARTICLE General Provisions 11170.1 Citation of Chapters 3 5 and 6 11170.2 Conduct of proceedings under Administrative Procedure Act 11172.3 Facilitate and support participation in hearing...ARTICLE 2 Administration of the Code 11191.1 Definitions of words and phrases used in this chapter...ARTICLE 3 Administrative Procedure Act Exemptions 11192.1 Definitions of words and phrases used in this chapter...ARTICLE 4 Administrative Procedure Act Exemptions 11193.1 Definitions of words and phrases used in this chapter...ARTICLE 5 Administrative Procedure Act Exemptions 11194.1 Definitions of words and phrases used in this chapter...ARTICLE 6 Administrative Procedure Act Exemptions 11195.1 Definitions of words and phrases used in this chapter...

5. [G09-40-1163817] San Diego Superior Court - CCP Breast Implant - Document text frame
   ...NOTICE OF COORDINATION AND vs ADOPTION OF MASTER COMPLAINT Defendants PLAINTIFF COMPLAINTS OF THE DEFENDANTS AND EACH OF THEM AS FOLLOWS!
   Plaintiff refers to and incorporates herein by reference a certain Master Complaint filed in IN RE COORDINATED BREAST... pursuant to Code of Civil Procedure Section 405.1. Pursuant to a coordination and inclusion in Judicial Council Coordination Proceeding No 254, pursuant to the Honorable Robert C. J. Ct. Judge of the Superior Court of the State of California for the... pursuant to Code of Civil Procedure Section 405.4...
Research Question I – Human vs. Machine

- Result

<table>
<thead>
<tr>
<th></th>
<th>Best Match (L)</th>
<th>Topic Distillation (L)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Auto</td>
<td>Human</td>
</tr>
<tr>
<td>Relevance</td>
<td>3.27</td>
<td>2.47</td>
</tr>
<tr>
<td></td>
<td>2.50</td>
<td>2.13</td>
</tr>
<tr>
<td>Depth</td>
<td>3.69</td>
<td>3.24</td>
</tr>
<tr>
<td></td>
<td>3.52</td>
<td>3.28</td>
</tr>
</tbody>
</table>

**Relevance**: The page is relevant to the topic
1=Agree strongly, 3=Neutral, 5=Disagree strongly

**Depth**: Is the page too broad, too narrow, or at the right level of detail for the topic?
1 = Too broad, 3=Right level, 5=Too narrow
Research Question I – Human vs. Machine

- Result: human + machine >> machine

\[ p < 0.001 \]

<table>
<thead>
<tr>
<th></th>
<th>Best Match (L)</th>
<th>Topic Distillation (L)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Auto</td>
<td>Human</td>
</tr>
<tr>
<td>Relevance</td>
<td>3.27</td>
<td>2.47</td>
</tr>
<tr>
<td></td>
<td>2.50</td>
<td>2.13</td>
</tr>
<tr>
<td>Depth</td>
<td>3.69</td>
<td>3.24</td>
</tr>
<tr>
<td></td>
<td>3.52</td>
<td>3.28</td>
</tr>
</tbody>
</table>

Relevance: The page is relevant to the topic
1=Agree strongly, 3=Neutral, 5=Disagree strongly

Depth: Is the page too broad, too narrow, or at the right level of detail for the topic?
1 = Too broad, 3=Right level, 5=Too narrow
Research Question I – Human vs. Machine

- Result: human + machine >> machine

\[ p < 0.001 \]

<table>
<thead>
<tr>
<th></th>
<th>Best Match (L)</th>
<th>Topic Distillation (L)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Auto</td>
<td>Human</td>
</tr>
<tr>
<td>Relevance</td>
<td>3.27</td>
<td>2.47</td>
</tr>
<tr>
<td></td>
<td>Auto</td>
<td>Human</td>
</tr>
<tr>
<td></td>
<td>2.50</td>
<td>2.13</td>
</tr>
<tr>
<td>Depth</td>
<td>3.69</td>
<td>3.24</td>
</tr>
<tr>
<td></td>
<td>3.52</td>
<td>3.28</td>
</tr>
</tbody>
</table>

**Relevance:** The page is relevant to the topic
1=Agree strongly, 3=Neutral, 5=Disagree strongly

**Depth:** Is the page too broad, too narrow, or at the right level of detail for the topic?
1 = Too broad, 3=Right level, 5=Too narrow
Result: A system without human interaction, optimized for topic distillation, performs as well as a non-optimized system with human interaction

<table>
<thead>
<tr>
<th></th>
<th>Best Match (L)</th>
<th>Topic Distillation (L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auto</td>
<td>Human</td>
<td>Auto</td>
</tr>
<tr>
<td>Relevance</td>
<td>3.27</td>
<td>2.47</td>
</tr>
<tr>
<td>Depth</td>
<td>3.69</td>
<td>3.24</td>
</tr>
</tbody>
</table>

Relevance: The page is relevant to the topic
1=Agree strongly, 3=Neutral, 5=Disagree strongly

Depth: Is the page too broad, too narrow, or at the right level of detail for the topic?
1 = Too broad, 3=Right level, 5=Too narrow
Research Question II
– Machine helps human

Hypothesis

- The performance gain from a topic distillation engine over a generic content-oriented search engine would be transferred to the human performance with each engine.
Research Question II – Machine helps human

- Result: Topic distillation is better than automatic best match.

<table>
<thead>
<tr>
<th></th>
<th>Best Match (L)</th>
<th>Topic Distillation (L)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Auto</td>
<td>Human</td>
</tr>
<tr>
<td>Relevance (1)</td>
<td>3.27</td>
<td>2.47</td>
</tr>
<tr>
<td>Depth (3)</td>
<td>3.69</td>
<td>3.24</td>
</tr>
</tbody>
</table>

Relevance: The page is relevant to the topic
1=Agree strongly, 3=Neutral, 5=Disagree strongly

Depth: Is the page too broad, too narrow, or at the right level of detail for the topic?
1 = Too broad, 3=Right level, 5=Too narrow

$\rho < 0.004$
Research Question II – Machine helps human

- Result: Performance improvement at system level transfers to interactive retrieval.

<table>
<thead>
<tr>
<th></th>
<th>Best Match (L)</th>
<th>Topic Distillation (L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auto</td>
<td>3.27</td>
<td>2.50</td>
</tr>
<tr>
<td>Human</td>
<td>2.47</td>
<td>2.13</td>
</tr>
<tr>
<td>Relevance (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Auto</td>
<td>3.69</td>
<td>3.52</td>
</tr>
<tr>
<td>Human</td>
<td>3.24</td>
<td>3.28</td>
</tr>
<tr>
<td>Depth (3)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Relevance: The page is relevant to the topic
1=Agree strongly, 3=Neutral, 5=Disagree strongly

Depth: Is the page too broad, too narrow, or at the right level of detail for the topic?
1 = Too broad, 3=Right level, 5=Too narrow

\[ p < 0.004 \]
\[ p < 0.01 \]
Hypothesis

- The interface for constructing a resource page would be more useful to the potential user if it reflects the structure of the problem domain, rather than simply presenting a list of unrelated documents.
Research Question III – Linear vs. Hierarchical

Methods

- Categorise/group search results according to their organisational/business units structure. (CSIRO + Rutgers)
- Summarise the information from each site with the title of the top three relevant pages. (CSIRO)
Interfaces from Rutgers – Hierarchical Interface

The best current approach ... per one or picker harvested cotton would be adequate to bring down virtually any new rate to an acceptable moisture content to initiate composting. Two systems were operated successfully to inject surfactant into water used to wet COMB or... from this source... cotton... can be done for greater underapplication rate for... Cotton... years earlier.

http://www.lib.ars.usda.gov/cppru/

New Items

http://www.lib.ars.usda.gov/cppru/cotton.htm

Cotton Production and Harvesting Research Selected Results Cotton Production Research Desiccation and Defoliation Harvesting Research Selected Publications General Information Cotton is one of the few plants which produces a spinable fiber. Cotton when grown in tropical climates is a perennial... Of the United States... where cotton is commercially grown... The seed are planted during the first two weeks of May... yield... fiber quality... length... strength... and maturity... and disease resistance... are the primary factors used to select cultivars... compared... and these were... recommended that cotton... should not be planted... the soil temperature... than 50 degrees F. There is...
Interaces from CSIRO – Site Summary

[Info: The top 100 retrieved entry points are grouped under their corresponding departments. Under each department, a maximum of three potential entry points are provided for each website. Clicking on any of these entry points will lead you to the corresponding website.]

1. United States Department of State
   - U.S. Embassy in Nicaragua - International Adoption, Nicaragua (usembassy.state.gov/managua/www/adopt.html) [Summary]
   - AMERICAN CITIZEN SERVICES 2090 (usembassy.state.gov/kansawwwwh2090.html) [Summary]
   - International Adoptions (travel.state.gov/intl/adoptions.html) [Summary]
   - Egypt - International Adoption (travel.state.gov/adoption_egyp.html) [Summary]
   - Nicaragua International Adoption (travel.state.gov/adoption_nicaragua.html) [Summary]

2. State of Louisiana
   - or orders or any Administrative Procedure Act RS 49 250 et seq fees Editors Note The following Act
     (www.lacourts.la.gov/oss/apa.pdf) [Summary]

3. Administrative Office of the U.S. Courts
   - FEDERAL RULES OF CIVIL PROCEDURE WITH FORMS DECEMBER 1, 2000 E PLURIBUS UNUM S Printed for the use of THE
     (www.ustcourts.gov/docs/civproc.pdf) [Summary]
   - RULE 1 1 RULES OF PROCEDURE SCOPE OF THESE RULES a The following Rules supplement the Federal Rules of Civil
     (pacer nd uscourts.gov/jfr/rules/rules.htm) [Summary]
   - Rule 1 18 -- Procedure for Enacting Additions or Amendments to the Rules of Practice for the District of Arizona
     (www.azd.uscourts.gov/azd/courtsinfo/1999/5214f1f2996dj60726603005907c3129234261276e3d30725664007aa913%
     30OpenDocument) [Summary]
   - Rules of Practice AND PROCEDURE United States Court of Appeals For The Armed Forces Effective February 27, 1996 As amended
     (www.armfor.uscourts.gov/rules.doc) [Summary]
   - amended Local Rule of Civil Procedure 53.2 3 effective September 13, 1999 These amendments will have the effect of
     (www.paeid.uscourts.gov/local/rules/ARBNOTr.pdf) [Summary]
   - CASE MANAGEMENT ORDER No 17 ADOPTION OF BRIGHT LINE TEST AND PROCEDURES With the assistance of Plaintiffs
     Lead Counsel end (www.paeid.uscourts.gov/documents/MDL/MDL1148/CMO17.HTM) [Summary]
   - CASE MANAGEMENT ORDER No 38 ADOPTION OF TEST AND PROCEDURES FOR DISMISSAL BASED ON STATUTE OF
     LIMITATIONS With the assistance (www.paeid.uscourts.gov/documents/MDL/MDL1148/CMO38.HTM) [Summary]
**Hypothesis III – Linear vs. Hierarchical (1)**

- Result – no significant differences

<table>
<thead>
<tr>
<th></th>
<th>Best Match</th>
<th></th>
<th>Topic Distillation</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Linear</td>
<td>Hierarchical</td>
<td>Linear</td>
<td>Hierarchical</td>
</tr>
<tr>
<td>Relevance</td>
<td>2.47</td>
<td>2.57</td>
<td>2.13</td>
<td>2.40</td>
</tr>
<tr>
<td>Depth</td>
<td>3.24</td>
<td>3.31</td>
<td>3.28</td>
<td>3.23</td>
</tr>
<tr>
<td>Coverage</td>
<td>2.67</td>
<td>2.92</td>
<td>2.48</td>
<td>2.53</td>
</tr>
<tr>
<td>Repetition</td>
<td>2.59</td>
<td>2.34</td>
<td>2.38</td>
<td>2.42</td>
</tr>
</tbody>
</table>
Hypothesis III – Linear vs. Hierarchical (2)

- Result – the hierarchic interface is conducive to less interaction

<table>
<thead>
<tr>
<th></th>
<th>Best Match</th>
<th>Topic Distillation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L</td>
<td>H</td>
</tr>
<tr>
<td>Iterations/queries</td>
<td>4.19</td>
<td>3.61</td>
</tr>
<tr>
<td>Time (seconds)</td>
<td>618.5</td>
<td>575.8</td>
</tr>
<tr>
<td>Number of surrogate seen</td>
<td>42.78</td>
<td>41.91</td>
</tr>
<tr>
<td>Number viewed</td>
<td>11.81</td>
<td>11.50</td>
</tr>
<tr>
<td>Number selected</td>
<td>10.64</td>
<td>10.25</td>
</tr>
<tr>
<td>Number ever saved</td>
<td>6.14</td>
<td>5.78</td>
</tr>
<tr>
<td>Number of final saved</td>
<td>6.00</td>
<td>5.63</td>
</tr>
</tbody>
</table>

The table above shows the comparison between linear (L) and hierarchical (H) interfaces for various metrics, with the hierarchical interface generally performing better in terms of interactions.
**Hypothesis III – Linear vs. Hierarchical (3)**

- Result – no significant difference in subjective self-assessment of task success

<table>
<thead>
<tr>
<th>Topic</th>
<th>Best Match</th>
<th>Topic Distillation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L</td>
<td>H</td>
</tr>
<tr>
<td>Perceived usefulness (7)</td>
<td>5.09</td>
<td>4.88</td>
</tr>
<tr>
<td>Perceived focus (7)</td>
<td>4.95</td>
<td>4.95</td>
</tr>
<tr>
<td>Perceived coverage (7)</td>
<td>4.63</td>
<td>4.58</td>
</tr>
<tr>
<td>Enough time (7)</td>
<td>4.75</td>
<td>4.38</td>
</tr>
</tbody>
</table>

All variables are on a 1..7 Likert scale (least to most).
Hypothesis III – Linear vs. Hierarchical (4)

- Result – consistent and sometimes significant preference for hierarchic interface

<table>
<thead>
<tr>
<th></th>
<th>Best Match</th>
<th>Topic Distillation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L</td>
<td>H</td>
</tr>
<tr>
<td>Easy to use</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>Task support</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>Preference</td>
<td>2</td>
<td>10</td>
</tr>
</tbody>
</table>
The automatic topic distillation engine did as well as the human users using a non-tuned search system.

The engagement of the user’s effort has a significant effect on the task performance.

The users performed significantly better with the search engine that is tuned to the topic distillation task.

For the comparison of the hierarchic delivery and the ranked list:

- There is no significant difference between the ranked list interfaces and the hierarchical interfaces in terms of the four performance criteria.
- The hierarchic interface is conducive to less interaction.
- Subjects preferred the hierarchical interfaces.
Conclusions

- Tuning the system to the task is a good idea.
- The structured display is a good idea.
  - For the same level of performance users made less effort.
  - Users preferred a display of the results that is tuned to the task.
Thank you ...