LABORATORY: Comparing Search Engines/Document Analysis

1. How do different search engines handle queries about the same topic? What are their respective strengths and weaknesses?

Working in pairs,

Go back to the quiz questions you explored via Google in an earlier, as well as the ‘facts’ about A.A. Milne provided in Lecture 10.

Try answering each of the questions in turn by using the following search tools:

- Yahoo Directory (NOT the search engine at the Yahoo home page) – www.dir.yahoo.com
- Ask Jeeves (www.ask.com)
- Mootersearch (http://www.mooter.com/moot)

How does each search tool rate in terms of:

- Recall and precision
- Authority of information
- Ease of searching technique
- Breadth of web coverage
- Help facilities

Are there any circumstances in which you would choose one of these tools instead of Google? Why?

2. Conduct your own search for web-based material comparing the strengths and weaknesses of different search engines. Some leads are provided in the further reading from last week’s lectures.

- What (if any) surprises have you encountered?
- Is there a consensus amongst commentators as to the relative strengths and weaknesses of search engines?
• How do the views you have read compare with your own findings earlier in
the lab?

Be prepared to spend the first 10 minutes of the tutorial session discussing these
findings.

**TUTORIAL: Document Analysis**

Document analysis is the unavoidable starting point for any classification process. We
can’t classify a document if we don’t know what to call it. Giving a document a label
implies that it has some commonly recognised characteristics or elements. Many of
these elements (not just “content”) contain the “meaning” of a document. In electronic
information systems it is particularly important that these elements be identified *and*
captured if some or all of its information value is not to be lost.

**In Pairs:**

Using the procedure outlined in last week’s lecture, attempt an analysis of the
documents provided, some of which come from recent exam papers. Use the attached
table to make notes on the documents under the following headings:

- Form (i.e. the **label** we attach to this type of document)
- Format (a “shape” given to information which makes it recognisable. I.e. how
  is the information structured in familiar and recognisable ways?)
- Medium (in or on which information is created/stored/accessed)
- Context (What do we need to know about this documentary form and its
  content before we can fully understand what it is. Imagine you are a “Martian”
  who has never seen one before. What would you need to know about it.)
- Authority (what “power”, if any, does it have? How do we know it is what it
  purports to be?)
- Content (What is it about? What does it contain?)
- Purpose (Why was it created? Kept? Etc.)
- Technology / technologies (used to create *and* access it)
- Accessibility (in what way do the other elements affect our ability to readily
  use it?)

**Document 3:**
Principal cardholder details

Full name: Frederick Forseythe
Address: 29 The Avenue, New
Telephone no. 9247-5432

Credit card account number: 5372-8346-834

Disputed transaction details

I wish to dispute the following transaction(s) recorded on my credit card statement:

Transaction details

<table>
<thead>
<tr>
<th>Date</th>
<th>Merchant</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>01/05/02</td>
<td>Ticketel</td>
<td>$55.00</td>
</tr>
<tr>
<td>27/06/02</td>
<td>Myers</td>
<td>$39.00</td>
</tr>
</tbody>
</table>

Tick the appropriate box and detail any additional information in the area provided below:

D I did not authorise or participate in the transaction(s) (nor did the additional cardholder). I do have all card(s) in my possession. D I did not authorise or participate in any Internet transaction(s) (nor did the additional cardholder).

D I only authorised one transaction (apparent duplication).

D I did engage in the transaction(s) but did not receive the goods/services ordered (mail/telephone order).

Expected date of delivery was:
I contacted the merchant on: ____________ and the merchant's response was:

D Transaction(s) incorrectly processed. I authorised only $ (copy of valid transaction(s) attached).
D Credit note has not been processed or has been listed as a charge on my statement (copy of credit note attached).
D I have cancelled my arrangement for a periodic debit with the merchant on: ____________ (copy of the cancellation letter must be provided).
D Other: give details

Additional information

Note: Should the transaction(s) in dispute be found to be legitimately authorised by yourself or any other authorised party, a voucher retrieval fee per disputed transaction may be applied.

Principal cardholder signature (mandatory): Frederick Forseythe
Additional cardholder signature: ____________
Date: 10/10/02

Important information for Cardholder:

Attach a copy of vouchers, statements and any documentation available that may assist in our investigation.
Retain a copy of this form and original documentation.
<table>
<thead>
<tr>
<th>Element</th>
<th>Document 1</th>
<th>Document 2</th>
<th>Document 3</th>
<th>Document 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Form</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Format</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Context</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Authority</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Content</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Purpose</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Technology/ie s</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Element</td>
<td>Document 5</td>
<td>Document 6</td>
<td>Document 7</td>
<td>Document 8</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>Form</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Format</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Context</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Authority</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Content</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Purpose</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Technology/ies</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
</tbody>
</table>