Turning Search into Find

Matthew Ellison
matthew.ellison@uaeurope.com

What we’ll cover in this session

- Why search is important
- The obstacles to Find
- Innovative approaches to search on the Web
- Top 10 factors that help turn Search into Find
- Some practical pointers towards implementation
Why search/find is so important

- Search is not necessarily the most effective tool for finding information
  BUT
- It is the tool that users prefer to use
  [Hood, Henderson, Jordan, 2003]
- Many Help systems now omit an Index

We used to find things...

- Now we search...
What are the obstacles to Find?

- Don’t know what search keywords to use
- Can’t ask questions
- Too many search results
- Search does no linguistic analysis
- Search is not picking up synonyms
  - Can’t specify search tightly enough
  - Zero search results
- Don’t really know what I’m looking for
- Search takes no account of semantics
- Not clear which search result is best
- Search has no understanding of the domain

Best search result is at the bottom of the list

SOME INNOVATIVE APPROACHES TO SEARCH ON THE WEB
Turning Search into Find

Google Suggest
(Predictive Search)
Turning Search into Find
Remind you of anything?

Google toolbar

Previously typed keyword
Turning Search into Find

Google toolbar update

![Google toolbar update](image)

No History section

BBC I-Player

![BBC I-Player](image)

Update:
- Seven Ages of Britain
- I'm Sorry, I'll Read That
- 6 Music Plays It Again
Turning Search into Find

Confused.com

Car Insurance Quotation for Matthew Ellison

Drivers

Your Occupation

What is your employment status? Director

What is your main occupation? Computer Analyst

Do you have another occupation? Yes

“Search” results

Railsaver.co.uk

RAILsaver

Journey type: One way

Passengers:

Adult: 1

Child (0-15): 0

Travelling from:

Rib

Travelling to:

Ribblehead

Departing: Wed

Leaving after: 18:00

Returning: Wed

Leaving after: 18:00

Are the 2nd and 3rd suggestions useful?
Imagine I’m searching for Waterloo (London). Required station included.
**British Airways (BA.com)**

- Simple Search
  - To: Where we fly
  - swed
  - Gothenburg, Sweden, GOT, Gothenburg
  - Stockholm, Sweden, STO, Stockholm (All)

- Advanced Search
  - To: Where we fly
  - stc
  - Stockholm, Sweden, STO, Stockholm (All)
  - arl
  - Stockholm, Sweden, ARN, Arlanda

---

**Need to Balance search functionality and simplicity/ease of use**

- HTML Help – Simple Search
  - Should we ever require users to understand Boolean logic?

- HTML Help – Advanced Search
Need to Balance search functionality and simplicity/ease of use

- Two online bookshops:
  - Borders (UK) – popular high street bookseller
    - Now out of business
  - Blackwells – strong links to academic institutions and libraries

- Compare their approach to search

Borders

- Predictive search
- Simple UI
- Can enter title, ISBN, or author
- Matches within titles
- Two result categories: products and people

World Champion Speed Reader, Anne Jones races through the Dan Brown's new novel in under 42 mins...

But what did she make of the much-anticipated follow-up to The Da Vinci Code? She described the book as having "all the usual Brown fare: clues and codes, mysteries and conspiracies, a race against time...and a gruesome
Turning Search into Find

Borders

Blackwells
Faceted Search

- Classify information by specific characteristics (facets)
- More flexible than a TOC, which presents a single, pre-determined, taxonomic structure
- Users explore available information by choosing required facets
  - Combine facets to narrow down the search
- Can be used to refine the results of full-text search
Example of faceted search

Browse Varietal
- Red Wines (171)
- White Wines (149)
- Bubbly (46)
- Pink Wines (139)
- Dessert/Port/Malvasia wines (41)

Browse Region
- French (55)
- German (5)
- South American (4)
- Other European (8)
- Portuguese (15)
- Spanish (15)
- USA (265)

Browse Price
- Bargains under $20 (227)
- Top shelf (over $100) (11)
- Set your own Price:

http://facetmap.com/browse/

Combining FTS and faceted search

[Image of a website showing a faceted search interface with options for gender, sport, and price, and a product listing with filters for saucy grid type and price range.]
FACTORS THAT HELP TURN SEARCH INTO FIND

In reverse order...

#10 Stop words – facility to exclude specific words from search matching

- Common words that are not indexed by search
- Marginally reduces file size and increases search speed
- Might cause problems when searching for phrases
  - Example: sort by date
#9 Facility to exclude specific topics from search

- Option to exclude topic from search
- Enables you to ensure topic only appears in a specific context
  - Context-Sensitive Help topics
  - Pop-ups
  - Topics with pre-requisites

Searchable:
- Include topic when full-text search database is generated

This feature is available for DotNet Help, WebHelp, and WebHelp Plus targets only.

#8 Search result synopses (context)

- Show first few words or key extract for each search result
- Similar to Google's presentation of search results
- Enables user to assess relevance of search result
#7 Boolean search

- **AND, OR, NOT**
- Enables users to combine keywords for advanced search
  - Using AND decreases number of results
  - Using OR increases number of results
- Most Web search engines use Boolean **AND**
- Most Help search engines use Boolean **OR**
  (perhaps afraid to offer zero results)

### Comparing search provided by HATs

<table>
<thead>
<tr>
<th>Help format</th>
<th>Boolean Ops</th>
<th>Other features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adobe AIR Help and WebHelp</td>
<td>Supported</td>
<td>Ranking</td>
</tr>
<tr>
<td></td>
<td>Defaults to OR</td>
<td>Highlighting</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Context</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Substring</td>
</tr>
<tr>
<td>MadCap WebHelp</td>
<td>Supported</td>
<td>Ranking</td>
</tr>
<tr>
<td></td>
<td>Defaults to OR</td>
<td>Highlighting</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Filtering</td>
</tr>
<tr>
<td>ComponentOne NetHelp</td>
<td>Not supported</td>
<td>Ranking (no numbers)</td>
</tr>
<tr>
<td></td>
<td>Defaults to OR</td>
<td>Highlighting</td>
</tr>
<tr>
<td>Author-it Browser-based Help</td>
<td>Supported</td>
<td>Ranking + hit count</td>
</tr>
<tr>
<td></td>
<td>Defaults to AND</td>
<td>Match partial words</td>
</tr>
<tr>
<td>WebWorks Help</td>
<td>Not supported</td>
<td>Ranking</td>
</tr>
<tr>
<td></td>
<td>Defaults to AND</td>
<td>(arbitrary numbering)</td>
</tr>
</tbody>
</table>
#6 Phrase matching (quoted keywords)

- Examples:
  - “currency conversion”
  - “topic templates”
  - “content explorer”

- Usually unnecessary with sophisticated search engines

- May help to filter out unwanted search results

#5 Fuzzy matching linguistic analysis

- Suggesting misspellings
- Offering close matches
- Stemming and parts of speech
  - Cow finds cattle
  - Translate finds translation and
#4 Search filtering and faceted search

- Concept of “Information Types” coined by Microsoft in 1996 for HTML Help
- Commonly used on e-commerce sites
- Achievable with Concept Keywords and Search Filters in Flare’s WebHelp and DotNetHelp

Facets for software user assistance

**Sports Shoes**
- Gender
  - Men
  - Women
- Sport
  - Running
  - Gym...
- Size
  - 4, 5, 6, 7, 8
- Type
  - Cushioned
  - Trail

**Software user assistance**
- Role
  - Administrator
  - User
- Software module
  - Accounts
  - Human resources
- Experience
  - Novice, intermediate, advanced
- Required info
  - Step-by-step
  - Field descriptions...
#3 Ranking of search results

- What determines sequence?
  - Number of occurrences of keyword
  - Meta data
  - Concept of “best bet” in MS Vista Help

#2 Meta Data

- The key to flexible and effective search
- Enables you to avoid zero search results
  - Topics found even if they don’t contain the search keywords within the visible text
- RoboHelp 8:
  - option to add search keywords manually
  - automatically adds index keywords as search keywords
  (you must not enable substring search)
#1 Predictive search

- Reduces required number of key presses
- Provides better “scent of information”
- Discourages “long tail” searches?
  - Before Google Suggest, 20 - 25% of all searches never occurred before
    (http://preview.tinyurl.com/longtail25)

Practical ways to implement predictive search

- Google Custom Search
  - Tips and guidelines available on the Web
  - Use PHP and jQuery to add auto complete (see http://tinyurl.com/nkfw8p)
- Turnkey auto-suggest technologies such as PredictAd
- Adobe Forums case study
Turning Search into Find

### PredictAd

- Learns community search patterns
- Provides:
  - suggested search queries (like Google Suggest)
  - direct links to adverts, live search results, related articles, etc. (like Borders Search)
- These results can be based on:
  - partial keyword entry
  - deep analysis of content
  - the user’s profile
  - the context of the search

---

### Adobe Forums

![Adobe Forums](image)

- **Information type**: Partial Search query
- **Search results** (not suggested keywords)
Adobe Forums

- Powered by Jive Software’s Clearspace
- How the predictive search works:
  - JavaScript captures keyup events from the search box
  - after a specific number of characters, an AJAX request is sent to the server with a wildcard search based on the characters entered for example: templ*
  - JavaScript is used to create an HTML div containing HTML that formats the results
  - Other JavaScript code enables the user to use the up and down keys to move through results and the enter key to select one and go to the content

Adobe Forums - update

Search query – No Auto-suggest!

Interesting “Related Topics” feature
References

- Results and Implications of Testing Search vs. Index for Lookup Tasks: Hood, Henderson, Jordan, 2003
- Mozilla Developer Center: How to implement a custom autocomplete search component http://tinyurl.com/mkkysw
- Predictive Search Query Suggestions http://www.seobythesea.com/?p=1375
- Drupal: Keyword Autocomplete http://drupal.org/project/keyword_autocomplete

Questions?

Matthew Ellison
matthew.ellison@uaeurope.com