

Legal Aid Drupal Users Collective

Collaboratively building open source expertise and innovations in legal services

drupal.openadvocate.org



Developing Website Search Systems: 2018 Research

What we need to know

Do we need a search system?

This will depend on:

- Volume of content
- Time required to set up and maintain a search system
- Payoff it will bring to the site

Results from recent studies show that:

- On average about 14% of participants started with the search (between 11% and 21% of actual users).
- There is no evidence of search-dominant users. Based on a recent study on e-commerce sites conducted by UIE (User Interface Engineering), none of the participants in the study used search exclusively but there were some link-dominant users (20% of participants). In the same study, 21% of the sites that UIE examined individually, every single user who visited the sites, used search. It seems that these sites were search dominant, not the users. In contrast, 32% of the sites, were link dominant (users used only the links on the site) and 47% of the sites were not search or link dominant.
- One factor that predicts whether users would initially start with search or with links is the type of product being sold on the site (certain types lend themselves better to being searched).
- Users also gravitate to the search engine when the links on the page don't satisfy them in some way.

The takeaways:

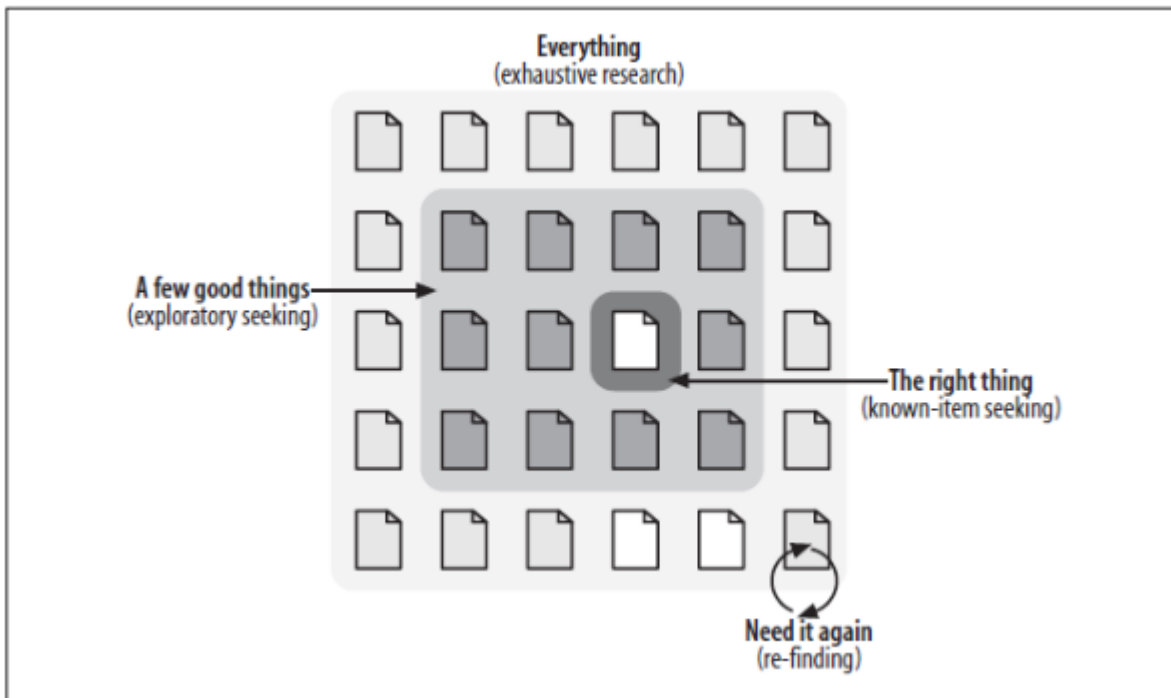
- A user's decision to search or navigate depends on
 - whether certain types of information lend themselves better to being searched
 - the website itself (e.g. inherent shortcomings)

- the ability of the user to predict a suitable path through the website
- Other things to consider
 - Search can be a learning tool - can tell us useful data on what users want from our site
 - Should be there because users expect it to be there
- Have a good search to support a good navigation.

Information Needs: What do users want?

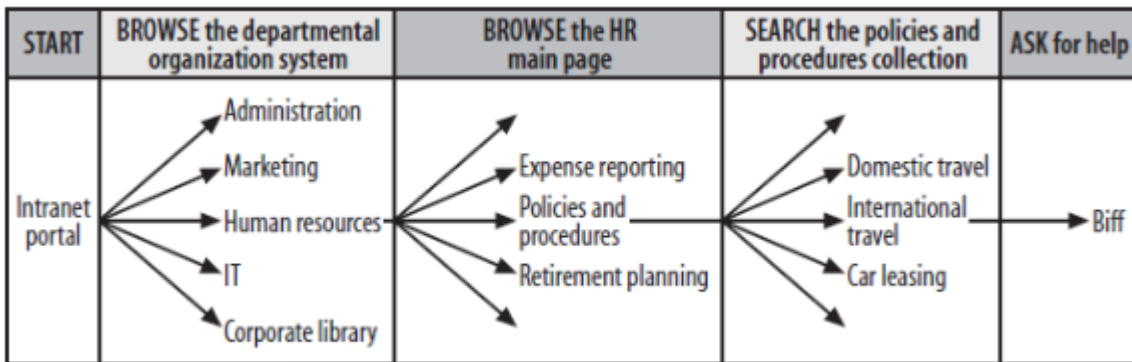
The 4 Common Information Needs (Fishing Metaphor)

1. The right thing (known-item seeking): The Perfect Catch
2. A few good things (exploratory seeking): Lobster Trapping
3. Everything (exhaustive search): Indiscriminate Driftnetting
4. Need it again (re-finding): I've seen you before, Moby Dick...



Two other major aspects to information seeking behavior

1. Integration - searching, browsing asking in the same finding session
2. Iteration - starts with information need, followed by formulating an information request (query), then moving iteratively through an information system along potentially complex paths, picking bits of information ("berries") along the way, modifying information requests in the process as users learn more about what they need and what information is available from the system



Another useful approach

The Pearl Growing Approach - users start with one or a few good documents that are exactly what they need (e.g. "more like this", "similar pages", etc.)

Other information seeking behavior concepts

- Reducing Uncertainty
 - Problem-solving
 - Identifying issues
 - Setting goals
 - Designing suitable courses of action
 - Decision making
 - Evaluating
 - Choosing among alternatives
- Browsing
 - Informal, unplanned
 - Aimless vs. goal-related
 - Scanning
 - Serendipity
- Relevance
 - Requires context
 - 'Aboutness' (e.g. on the topic) vs. non-topicality
 - Pertinent - connected to a need
- Salience
 - Unexpected, notable, prominent
 - Selective exposure
- Avoiding information
 - Selective exposure (filtering)
 - Rejection of ideas
 - Reluctance to be distracted
 - Unused information

Factors that influence information seeking

- Information poverty
 - Knowledge gaps
 - Barriers to knowledge
 - Cultural patterns
 - Lack of processing skills (e.g. reading, language, hearing, sight)
- Information overload and anxiety. “We can seek knowledge in order to reduce anxiety and we can avoid knowing in order to reduce anxiety.” - Maslow
 - Omission
 - Error
 - Queuing
 - Filtering
 - Approximation
 - Multiple channels
 - Escaping

Information Seeking Theories

Principle of Least Effort

- An individual will adopt a course of action that will involve the expenditure of the probable least average of his work
- Asking nearest co-worker, using nearest tools, consulting older resource instead of a more current one, using interpersonal resources versus authoritative resources, relying on close relatives and close friends
- Cost-benefit paradigm
 - Trade-off between the effort required to employ a strategy and the quality of the resulting action
 - Seekers will minimize the effort required to obtain information even if it means accepting a lower quality or quantity of information

Uses and Gratification

- Audience plays an active (not passive) role in selecting sources - the person uses the medium, not the other way around
- Media are only a portion of a range of options for fulfilling needs
- Use can be studied by asking people directly
- How the poor deal with problems
 - Reasons for non-active information seeking

Play Theory and Entertainment Theory

- Pleasure principle - humans manipulate their intake of information to serve their emotional needs
-

Humans tend to:

- Seek pleasure and avoid pain
- Mix work with play

How we can learn more about our users' information needs and information seeking behaviors

1. Surveys
2. Interviews
3. Usability Study
4. Focus Group

Building the search system

Information Architecture: The Top-down Questions

1. Where am I?
2. I know what I'm looking for; how do I search for it?
3. How do I get around this site?
4. What's important and unique about this organization?
5. What's available on this site?
6. What's happening there?
7. Do they want my opinion about their site?
8. How can I contact a human?
9. What's their address?

Basic Anatomy

Search Aids

1. Search interface - means of entering and revising a search query with information on how to improve query as well as other ways to configure search
2. Query language - grammar of a search query
3. Query builders - ways of enhancing a query performance (e.g. spellcheckers, stemming tools, etc.)
4. Retrieval algorithms - determines which content matches a user's query
5. Search zones - subsets of site content that have been separately indexed to support narrower searching
6. Search results - presentation of content

"Invisible" Components

1. Controlled vocabularies and thesauri - predetermined vocabularies of preferred terms that describe a specific domain

2. Retrieval algorithms - used to rank search results by relevance
3. Best Bets - preferred search results that are manually coupled with a search query; editors and SMEs determine which queries should retrieve best bets, and which documents merit bet status

What to make searchable

1. Navigation vs. destination: When a user searches a site, they are looking for destination pages or pages that contain actual information; navigation pages get a user to the destination pages. If navigation pages are included in the retrieval process, they will just clutter up the results. This obstructs the user's ability to find the most useful result.
2. Indexing by search zones
 1. Content type
 2. Audience
 3. Role
 4. Subject/topic
 5. Geography
3. Indexing for specific audiences
4. Indexing by topic
5. Indexing recent content

Selecting content components to index

- Body
- Title
- URL
- Site name
- Image link
- Image alt text
- Description
- Keywords
- Remote anchor text

Retrieval Algorithms

Pattern Matching

- Recall and precision
 - Recall = # relevant documents retrieved/# relevant documents in collection
 - Precision = # relevant documents retrieved/# total documents in collection

Other Approaches

- Document Similarity - converts document into the equivalent of a query; "stop words" like "the" "is", "he" are stripped out of a good document, leaving useful set of semantically rich terms that, ideally, represent the document well, and converted into a query that should retrieve similar results

- Alternative Approach - present results that have been indexed with similar metadata

Query Builders

- Spell Checkers - allow users to misspell terms and retrieve the right results by automatically correcting search terms (e.g. accomodation, accommodation), ensuring retrieval of results that contain the correct term.
- Phonetic Tools - expands a query to include other possible results that matches the same phonetic sound (e.g. Smith, Smyth)
- Stemming Tools - allow users to enter a term and retrieve documents that contain variant terms with the same stem (e.g. parent, parenting)
- Natural Language Processing Tools - examines the syntactic nature of a query (Is it a “how to” question or “who is” question, and use that knowledge to narrow retrieval)
- Controlled Vocabularies and Thesauri - expand the semantic nature of a query by automatically including synonyms within the query

Presenting retrieval results

Present less to users who know what they are looking for and more to users who are not sure what they want. Example, display representational content components such as title or author to those who know what they want and descriptive content such as part of an abstract, summary or keywords to those who are not sure of what they want.

The Multiple Top Lists Presentational Model

- Gives the user an overview of what's available
- Grouped according to the following categories
 - Publications
 - Images
 - Maps
 - News
 - Projects
 - Tools
 - Staff
- Provides the top results for each type of document
- Guarantees user will get a sample from all document types on the page
- Identifies types of documents in the search results
- Trains the user as to what types of assets are available for searching
- Prevents each type of document from being swamped by results from other sources or types
- Provides “jumping-off” points for more detailed searches
- Organizes search results so that the eye is drawn to the most valuable information first

Other ways to present search results

- Filtered navigation
- Leave search term and let people edit it

- Incorporate forms of social proof into search results
- Let people sort results
- Provide viewing options (e.g. list view, gallery view)
- Anticipate spelling mistakes where possible and provide relevant results
- Avoid zero results or dead ends
 - Provide options
 - Advice/tips on searching
 - Provide search suggestions (e.g. show recent or most popular searches)

How many documents to display

- Let users know the total number of retrieved documents.
- Provide a results navigation system (pagination) to help users move through the result set.
- Listing results
- Sort by alphabet
- Sort by chronology
- Rank by relevance
- Rank by popularity
- Rank by user's or expert's ratings

Designing the search system

Search interface design

Consider

- Level of searching expertise and motivation
- Type of information needed
- Type of information being searched
- Amount of information being searched
- Dependent on the characteristics of the site and user

Test user assumptions

- I can just type terms that describe what I'm looking for and the search engine will do the rest
- I don't have to type in those funny AND, OR or NOT thingies.
- I don't have to worry about synonyms for my term; if I'm looking for dogs, I just type 'dogs,' not 'canine' or 'canines'.
- Fielded searching? I don't have time to learn which fields I can search.
- My query will search the entire site.

In the above scenarios

- Give users the box.
- Provide a help page that explains how to create more advanced, precise queries, but users may rarely visit this page.

- Too many or too few (zero) are both indicators for users to revise their searches.

Advanced Search

- Few users ever take advantage of them.
- It is something of a convention and users will expect to see it.
- Design search system with the goal of making it unnecessary for the vast majority of searchers to ever need to go to the advanced search page.

Helping the user hone search

- Repeat search in results. Displaying the initial search within the search box restates the search that was just executed, and allows the user to modify it without re-entering it.
- Explain where results come from.
 - Explain what the user did.
 - Restate the query
 - Describe what content was searched
 - Describe any filters that was used
 - Show implicit Boolean or other operators
 - Show other current settings such as sort order
 - Mention the number of results retrieved

Integrate search with browsing

- Making the search (and browse) clearly available to users by way of high contrast with the background and a prominent location on the page can preclude abandonment of a site if one method is not fruitful.

When users get stuck

- A means of revising search
- Search tips or advice on how to improve the search
- A means for browsing
- A human contact if searching and browsing won't work.

Some Examples

Presenting users with different ways to search - an example of reducing uncertainty and selective exposure

Helping users hone search with autocomplete

An example of search that applies the concept of integration and iteration (integrating search with browsing)

Sources

- Search Technologies. (2018). A Modern Search Results Presentation. [online] Available at: <https://www.searchtechnologies.com/blog/search-results-presentation> [Accessed 1 Feb. 2018].
- Amazon.com. (2018). Amazon.com: Online Shopping for Electronics, Apparel, Computers, Books, DVDs & more. [online] Available at: <https://www.amazon.com/> [Accessed 1 Feb. 2018].
- UX Articles by UIE. (2018). Are There Users Who Always Search?. [online] Available at: https://articles.uie.com/always_search/ [Accessed 1 Feb. 2018].
- Slideshare.net. (2018). Information Seeking Theories And Models. [online] Available at: <https://www.slideshare.net/guestab667e/information-seeking-theories-and-...> [Accessed 1 Feb. 2018].
- Marketing Strategies, D., Community, M., VRBlockchainChatbotsMobile, E., Hospitality, D., Charlton, G., Developments, D., Media, S., Marketing, C., Marketing, E., Advertising, D., Developments, I., Sales Content Collaboration Search, M., Analytics, D., Community, S., Advertising, D., Automation, E., Sales, E., VR Chatbots Mobile, E., VR, A., Technology, R., Marketing, D., Technology, M., Transformation, D., Technology, E., Technology Education, D., Leadership, S., Technology, T., Calculator, S., Calculator, C., Calculator, C., Calculator, R., us, A., newsletter, S., article, S., business, P., Policy, C. and Conditions, T. (2018). 17 ways to improve your site search results pages - ClickZ. [online] ClickZ. Available at: <https://www.clickz.com/17-ways-to-improve-your-site-search-results-pages...> [Accessed 1 Feb. 2018].
- Measuringu.com. (2018). MeasuringU: Search vs. Browse on Websites. [online] Available at: <https://measuringu.com/search-browse/> [Accessed 1 Feb. 2018].
- Morville, P. and Rosenfeld, L. (2007). Information architecture for the World Wide Web, 3rd ed. Sebastopol, Calif: O'Reilly.
- Ca.gov. (2018). Welcome to the California State Web Portal. [online] Available at: <http://www.ca.gov/> [Accessed 1 Feb. 2018].

This research was funded by a Legal Services Corporation Technology Initiative Grant (TIG) and compiled in June 2018 by Illinois Legal Aid Online and Pine Tree Legal Assistance.

Printed: October 17, 2019

<http://drupal.openadvocate.org/node/44/developing-website-search-systems-2018-research>

©Legal Aid Drupal Users Collective