

Making Search Work for You: Optimizing Content and Meeting Audience Expectations

Travis Hicks

American Society of Clinical Oncology

February 14, 2019

Agenda

1. Introduction
2. Optimizing Content for External Discoverability
 - Overview of benefits
 - 10 steps to help
3. Improving Internal Search
 - Understanding user base intent
 - Utilizing thesauri
 - Understand the technology
4. Case Studies
 - ASCO.org
 - ASCO Meeting Library
5. Conclusion

Introduction

Why Should You Want to Understand Search

- External search engines #1 mode for discovering unknown content
- But querying takes many forms
- To leverage, you want to account for all in
 - Constructing your content
 - Improving your internal search

Optimizing for External Discoverability

Good Content is Discoverable Content

- Google parses content not just for keywords, but also phrasing providing additional insight into relevancy
- Identifies terms that are semantically linked in its analysis of billions of documents
- In fact, Google has disincentivized poor content stuffed with keywords
- Content with a high-user value is discoverable content

10 Steps to Help Boost Your Search Results

1. Focus on your users and their intent in content
2. Write clean, concise copy
3. Create links to reliable, high-quality internal and external content
4. Use structured schemas that leverage updated html tags (schema.org)
5. Use bullets and organized lists
6. Ensure your content is mobile-friendly
7. Utilize taxonomies
8. Make sure site performance is strong—fast loads rank better
9. Ensure physical address is machine readable
10. Pay for it with Google Ad Words

Improving Internal Search

Better Understanding Base User Intent

- Review internal analytics
 - What types of queries do your users use (e.g., keywords vs. natural language)?
 - What are the topical buckets of your queries (e.g., what are users generally looking for)?
 - What terms are searched but don't produce any results?
 - Are your users more likely to search internally or externally?
 - If you index content from multiple sites, do users search to get to those sites?
 - If you have facets, do users actually use?
- Conduct user research
 - What is the overall level of satisfaction with search results?
 - If satisfaction is low, what trips them up (base usability can help determine)?
 - What areas or content types are the least discoverable?
 - What facets may be helpful in enabling the users to access content they are seeking?

Create and Support Thesauri

- A thesaurus is a hierarchical classification system of terms
- Thesauri employ synonyms, establish semantic relationships
- Adding terms as metadata enhance discoverability
- Synonyms can be utilized by search platforms like Solr
- Enables organizations to identify previously unknown relationships across content

Understand Possibilities/Limitations

- Your internal search isn't Google
- What can your internal search platform do?
 - Is there functionality you are not leveraging?
 - If there is, what do you need to unlock?
- What staff resources do you have available?
 - Do you have internal development staff?
 - Do you have individual(s) tasked with ongoing evaluation?
 - Do you have a search governance plan?
 - Is someone ultimately accountable for the search performance?

Case Study: ASCO.org

Overview of ASCO.org

- Overall organizational website
 - ❑ Provides information to members, public about ASCO
 - ❑ Serves as portal to other ASCO digital properties
- Overall site usage
 - ❑ 1.43M total visits
 - ❑ 4.5 million pages views annually
 - ❑ 894,000 total visits from external search
 - ❑ 118,000 unique searches executed in 2018
- Two primary user groups
 - ❑ Heavy users (board members, volunteers, internal staff)
 - ❑ Occasional users (average member, meeting attendees, pharma staff, public)

Current State and Challenges with Search

- ASCO.org employs Solr as its search engine
- Model and relevancy algorithm last overhauled in 2014
- Some ongoing maintenance, but no major work done since 2014
- Primary issues needing addressing:
 - ❑ Dissatisfaction by all user groups with relevancy of results
 - User expectations shifted with explosion in use of Google
 - Relevancy tuning has not kept up with content and business needs
 - ❑ User interface
 - ❑ New technology and functionalities

What Search Logs Told Us

- Primarily true keyword queries in Top 2,000 terms
 - 22 entries that could be considered natural language searches
 - Only 11 queries that could be considered advanced searches
- ASCO lacked content for some queries
- Other content needed to be refocused based on queries
- More than half of queries involved basic cancer types, treatments

Breakout Category Top 500 Terms

ASCO.org Top 500 Search Term Distribution Analysis: 2017-18			
Category	Search Terms Per Category	Searches Executed per Category	% Total Search Volume (Top 500)
ASCO meetings	23	3330	6.05%
ASCO publications	57	6000	10.91%
ASCO services/digital products	39	4639	8.43%
Basic cancer topics	112	17512	31.84%
Biology	17	1975	3.59%
Errors	2	131	0.24%
Genes	10	911	1.66%
Medicine	31	2871	5.22%
Member/attendee actions	28	3375	6.14%
Oncology professional organization	2	261	0.47%
Oncology professional practice	8	746	1.36%
Patient care	31	2830	5.15%
Policy	7	419	0.76%
Research study	9	460	0.84%
Treatment	149	13631	24.78%

Plan for Improvement

- Adjusting proximity parsing
- Limiting stemming
- Enhancing taxonomies
- Add new/improve search models
- Implementing relevancy score floor
- Improving the user interface

Proximity Parsing Adjustments

- Currently use AND to connect multiple word, which the two terms only to be used in the same document, not near one another
- Will move to a tiered approach with three initial buckets
 - ❑ Exact phrase match as entered
 - ❑ All words used within three words of one another
 - ❑ All words used within 20 words of one another
- Will test results to hone the buckets

Implementing Rules to Limit Stemming

- Stemming is the automatic stripping of a word to its word root, which has benefits but also unintended consequences
- For example, if a searcher looks for “Dues” (as in how much are dues or I want to pay my dues), Lucene will automatically remove the “s” and include “due” as well as “dues” in the result set
- Searcher then faced with results including “due to....”, “due on...” or that have a “due date”—none of which help a user looking to pay his or her member dues

Taxonomy Changes

- Utilize term weighting in relevancy calculations
- Expand existing organizational taxonomy to include member-based actions (e.g., pay dues)
- Create new authority file with board member names, VPs, other important individuals
 - This also would include multiple permutations of names as synonyms
- Bolster synonyms for existing terms based on search logs

Implementing Relevancy Score Floor

- Intended to reduce number of overall results (noise)
- If content doesn't reach a certain relevancy threshold, it will not be presented as a result
- Threshold will be determined based on testing after initial changes (e.g., where do results tail off?)

New Models/Adjustments to Existing Models

- Use new models to drive external content based on review of the search logs
- Add additional terms to existing models based on review of search logs

Construct of a Model

Name *

Label

Add a label for search result.

Description

The ASCO Membership Directory is your go-to tool for finding and interacting with other members of the American Society of Clinical Oncology (ASCO). Designed to keep you in touch with like-minded colleagues, the Membership Directory enables you to build your professional network, connect with colleagues at meetings and easily provide other ASCO members' information to patients seeking a referral.

Link

Boost

- None - ▾

Select the possible boost for this document. This will determine by how much the result will get pushed up in the results set.

Tags

membership directory, ASCO members, membership information, become a member, join ASCO, asco membership, find a member, member directory, members, member list, fasco members

Enter in a set of keywords or tags that can be used to promote the content within the search results. Please enter each separated by a comma.

UI Changes

- Add a type-ahead
- Contextual search
 - Did you mean?
- Distinguishing actions from normal results
 - Join ASCO
 - Pay dues

Measuring Success: Usability Testing

UX benchmarking provides feedback from end users on the search experience over time as improvements are implemented

- **Step 1: Baseline usability testing**

- Users complete search related tasks with search functionality and UI as-is.
- Documents usability of search/satisfaction with the current state

- **Step 2: Subsequent testing**

- As major milestones are met to improve search, users perform the same set of tasks completed during baseline testing
- By comparing results in each round of testing, we can confirm improvements have a positive impact on search usability and identify additional areas for improvement

- **Step 3: Rinse and Repeat**

Long-Term Possibilities

- Machine assisted relevancy optimization
 - Use SOLR logs and analytics to aid indexing and relevancy
 - Driven click click-through data
- Personalized search
 - Initially can be based off simple history, including profile data
 - Future would add known online user behavior and individual click data
- Voice assisted search

Case Study: Meeting Library

Differences in Users from ASCO.org

- Primarily meeting attendees and researchers (public, private)
- Users look to
 - Review presentations they attended
 - Catch up on presentations they missed due to schedule conflicts
 - Research concepts across meetings
- Craft more complex, targeted queries
- Utilize faceted filters
- Typically not members of the lay public

Differences in User Interface: Meeting Library

Search Results

SORT BY:

- Relevance
- Newest
- Oldest

FILTER BY SITE:

- ASCO.org (89)
- Meeting Library (4934)
- ASCO Connection (292)
- ASCO University (24)
- Meeting Sites (173)

FILTER BY TYPE:

- Abstracts, Presentations (4934)
- Guidelines (11)
- News (73)
- Web Pages (5)

Search returned 5023 results for "her2"

ASCO.org / GUIDELINES / November 14, 2016

HER2 Testing and Clinical Decision Making in Gastroesophageal Adenocarcinoma

This guideline provides specific recommendations for assessment of HER2 in patients with advanced gastroesophageal adenocarcinoma (GEA) while addressing pertinent technical issues and clinical implications of the results.

ASCO.org / GUIDELINES / September 2, 2014

Chemo- and Targeted Therapy for Women with HER2 Negative (or unknown) Advanced Breast Cancer

This guideline provides recommendations to identify optimal chemo- and targeted therapy for women with HER2 negative (or unknown) advanced breast cancer.

Differences in User Interface: ASCO.org

The screenshot shows the ASCO Meeting Library search results for the query 'her2'. The page features a search bar at the top with the text 'her2' and a search icon. Below the search bar is a navigation bar with filters: Filter, Year, Meeting & Track, Type, Topic, and Media. The results section shows 12651 results found for 'her2'. The results are displayed in a grid of six cards. Each card includes the meeting name, date, title, session information, author(s), presenter, and abstract number. The first three cards are from the GASTROINTESTINAL CANCERS SYMPOSIUM, and the last three are from the ASCO ANNUAL MEETING.

ASCO Meeting Library Sign In

her2

Advanced Search

Filter: Year Meeting & Track Type Topic Media

12651 results found for her2 Sort by View

GASTROINTESTINAL CANCERS SYMPOSIUM
Thursday, January 17, 2019
Comparison of HER2 expression by immunohistochemistry and HER2 copy number by digital PCR...
Session: Poster Session A: Cancers of the Esophagus and Stomach
Author(s): Naoki Takahashi, Satoru Iwasa, Tomohiro Matsu...
Presenter: Naoki Takahashi, PhD
Abstract #: 58

ASCO QUALITY CARE SYMPOSIUM
Friday, September 28, 2018
Education, engagement, and MIPS: The new era of accountability.
Session: Poster Session A: Big Data Studies; Projects Relating to Equity, Value, and Policy
Author(s): Diane Denny, DBA, Danielle Gross, Maurie Mark...
Presenter: Diane Denny, DBA
Abstract #: 111

GASTROINTESTINAL CANCERS SYMPOSIUM
Thursday, January 17, 2019
Trastuzumab beyond progression in patients with HER2-positive advanced gastric adenocarcinoma: A...
Session: Poster Session A: Cancers of the Esophagus and Stomach
Author(s): Yang Chen, Guanghai Dai; Chinese People's Libe...
Presenter: Yang Chen, M.D., Ph.D
Abstract #: 113

ASCO ANNUAL MEETING
Saturday, June 2, 2018
How to Improve Global Breast

ASCO ANNUAL MEETING
Monday, June 4, 2018
Panel Question and Answer

ASCO ANNUAL MEETING
Friday, June 1, 2018
Panel Question and Answer

Conclusion

Wrapping Up

1. Understand the expectations of your users
 - External searchers
 - Internal searchers
2. Develop a strategy for content optimization and stick to it
3. High-value user content is, by its very nature, discoverable
4. Not a silver bullet for improving internal search
5. All efforts need ongoing evaluation and iterative testing of improvements