Factors affecting visibility in web search engines

Jonas Fransson, PhD
E-resource Librarian at University Library
Visibility, findability and relevance

• How visible (highly ranked) web page in a web search engine depends on:

1. Content aspects of the page in relation to the search query -> **relevance**

2. Structural aspects of the page -> **findability**
Findability ≠ relevance

• Findability is query-independent
  – Not related to topic or search terms

• Relevance is query-dependent
  – Based on topicality or search terms
The findability-flower (Walter)

Morville’s definition

• Morville defines findability in *Ambient findability* as follows (2005, p.4):
  
a) The quality of being locatable and navigable.
  
b) The degree to which a particular object is easy to find or locate.
  
c) The degree to which a system or environment supports navigation and retrieval.
Two dimensions of web findability

• **Internal** web findability
  – how easy an object is to find within a web site.

• **External** web findability
  – how easy an object is to find on the web (outside the web site).
Findability factors on different levels

• On page factors
  – Type of object: web page, pdf, sound-file
  – Amount of text/metadata

• Site factors
  – Settings in the web publishing tool
  – Information architecture

• Web factors
  – Linking (inlinking)
Six aspects of findability

• Object attributes
• Accessibility
• Internal navigation
• Internal search
• Reachability
• Web prestige
Object attributes

• The content, form and format of the information objects. Important parts are metadata, file format and full text (if available).

• ≠ aboutness / topicality

---

**Discovery system i en ekologisk kontext**

*Jonas Fransson*

**Abstract**


Full Text: [PDF](http://rauli.cbs.dk/index.php/revy/article/view/4365)

Revy - ISSN: 1904-1969 - ISSN (online): 1904-1977

Hosted by [CBS Library](http://rauli.cbs.dk/index.php/revy/article/view/4365)
Accessibility

• The possibility to access information objects. It focuses on accessibility issues for persons with disabilities, but a disability-friendly website is also a search engine friendly site.

• New standard: Web Content Accessibility Guidelines (WCAG) 2.0 ISO/IEC 40500:2012

• Good: Plain text

• Bad: Scripts, text in pictures/videos, frames
Internal navigation

• A central part of an information system on the web. It is about link structures and the possibility to follow links down to individual objects.
Internal search

- The second way of navigation within websites. How well an internal search works depends on the search engine performance and settings together with the attributes of the objects.
Reachability

• Links to an object is a requirement for both users and search engines to reach the objects. Includes indexability by search engines.
Lund University Libraries

**eJournals & eBooks A-Z**  **Databases A-Z**  **LUBsearch**  **For Librarians**

| A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | R | S | T | U | V | W | Y | Z |

**General**
- Art, Music & Theatre
- Business & Economics
- Environment
- Humanities & Theology
- Law
- Mathematics
- Medicine & Health
- Natural Science
- Social Sciences
- Technology & Architecture

**Technology & Architecture**
- ACM Digital Library
- Art Index (EBSCOhost)
- Avery Index to Architectural Periodicals (EBSCOhost)
- BINGDOK: Kunskapsbanken, 1966 - 2006
- Civil Engineering Database (CEDB)
- DAAI: Design and Applied Arts Index
Web prestige

• The measure of inlinks to an object on the web. The prestige of the inlinks is an important factor beside the number of inlinks. The measurement can be seen as the possibility to land at the object at a random choice. An example is Google’s PageRank.
PageRank 1

PageRank 2

Measuring findability

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Measurement</th>
<th>Score</th>
<th>Total findability</th>
<th>External findability</th>
<th>Internal findability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Object attributes</td>
<td>Level of SAP (a)</td>
<td>0-3</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Full text (b)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accessibility</td>
<td>Compliance to WCAG-test</td>
<td>0-3</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Internal navigation</td>
<td>Object linked to in internal link navigation</td>
<td>0-1</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Internal search</td>
<td>Object indexed in internal search engine</td>
<td>0-1</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Reachability</td>
<td>Stable and unique URL</td>
<td>0-1</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Web prestige</td>
<td>PageRank value of the object</td>
<td>0-3</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

Maximum score 12 10 8

SEO and ASEO

• Search Engine Optimization
  – The practice of improving web pages for better visibility/ranking

• Academic Search Engine Optimization
  – How to improve academic documents for better visibility/ranking
## Comparison SEO/ASEO

<table>
<thead>
<tr>
<th></th>
<th>SEO</th>
<th>ASEO</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type of document</strong></td>
<td>Dynamic documents</td>
<td>Static documents</td>
</tr>
<tr>
<td><strong>Optimizing object</strong></td>
<td>Ongoing</td>
<td>Before publishing (by creator)</td>
</tr>
<tr>
<td><strong>Optimizing publishing</strong></td>
<td>Ongoing</td>
<td>Ongoing (by publisher)</td>
</tr>
<tr>
<td><strong>Prestige</strong></td>
<td>Inlinks (web prestige)</td>
<td>Citations (+altmetrics)</td>
</tr>
</tbody>
</table>
Academic findability – optimal process

1. Final version av article
2. Published (web access)
3. Publisher sends metadata to search services
4. Article indexed by bibliographic databases
5. Article findable in relevant databases and search services
Academic findability – possible issues 1

• Article has no standard URL or unique identifier (DOI)
  – Hard to link to through link resolver (LU Linker) based on metadata

• User has to accept agreement before entering publisher’s site
  – No direct linking
Academic findability – possible issues 2

• Publisher don’t deliver metadata to search services and link resolvers
  – Hard to find or link to article

• Journal not indexed in any major bibliographic database
  – Only publisher metadata available

• Article only published in full text, no or little metadata
  – Hard to find on just author or title
SEO for Google et al.

• Good, readable text
• Relevant inlinks and outlinks
• Descriptive anchor texts

• The text competes with all other texts that are relevant to the used search terms!
  – First relevance
  – Then ranking (incl. PageRank)
• Thank you for listening!

• Questions?

• Discussion?
Reading tips


