Content Management Systems (CMS)

Lecture 12: Search engines vs. Content Management (2)

Mariusz Trzaska, Ph. D.

mtrzaska@mtrzaska.com, http:// www.mtrzaska.com

Previous lecture continuation

Audit of an existing portal

- Usability,

- Availability for search engine robots,

- Checking currently indexed portal resources:
  - `Site:mydomain.com`,
  - Evaluation of results for keywords,
  - Check the version of the search engine cache (if it is current).

Audit of an existing portal (2)

- Verification of duplicate content
  - *mydomain.com* address should redirect(301) to www.mydomain.com/ (note the final slash).
  - As a result, the content of your site should appear in the SERP only for *www.mojadomena.com/*
  - You can use the `inurl` operator and `intitle` to see if we get only one result.
  - Make sure that specific content is available only under one URL.

Audit of an existing portal (3)

- Checking URLs
  - Keywords,
  - Minimum (zero) number of parameters
- Evaluation of page titles
  - Tag `title`,
  - Unique and descriptive,
  - Less than 70 characters,
  - The name of the organization should be at the end.

Audit of an existing portal (4)

- Content rating.
- Check meta tags
  - `robots`
  - `description` - used for page description in SERP.
- Files
  - *Sitemaps.xml,*
  - *Robots.txt.*
- Redirect (use code 301)

Audit of an existing portal (5)

- Linking inside the portal.
- Avoid unnecessary sub domains (cause "blur" of the main domain value).
- Geolocation - only relevant in some cases
  - geographic location of hosting servers,
  - information on the geographic location of the services „offered" by the portal, e.g. *microformat* `h-geo`.

Audit of an existing portal (6)

- External links to the portal
  - Tools
    - https://www.bing.com/toolbox/webmaster
    - https://moz.com/link-explorer
    - http://www.majesticseo.com/
  - It is important to take care of *deep linking* - links to the individual pages of the portal (e.g. a dedicated widget).
  - What does our link profile look like compared to the competition?

Audit of an existing portal (7)

- Time of loading.
- Utilization of the `alt` attribute in images.
- Accordance with standards - W3C validation.
- Analytical tools:
  - http://www.google.com/analytics/
  - https://www.bing.com/toolbox/webmaster
- Log files
  - https://www.graylog.org/products/open-source
  - http://www.logalyze.com/
  - https://www.nagios.com/
  - http://awstats.sourceforge.net/

Audit of an existing portal (8)

- Tools to work with search engines:
  - https://www.google.com/webmasters/tools/
  - http://www.bing.com/toolbox/webmasters/
  - https://webmaster.yandex.com/welcome/

Keyword Cannibalism

- It occurs when a term is used on many pages of the portal (mainly in the title).
- Avoid this situation, because the search engine must choose (how ?!) the one right page.
- In addition, such a reduced weight is less significant.
- In such a situation:
  - We choose the most important one,
  - And link other pages to it.

Hosting servers, and SEO

- This is not a key issue, but it is worth considering:
  - Server *timeout* - no response at specified time.
  - Long response time (page generation).
  - Shared IP addresses (potentially a bad reputation from previous users).
  - Blocked entire IP address pools.

Hosting servers and SEO (2)

- This is not a key issue, but it is worth considering - *continued*:
  - Limitations on the number of downloads per unit time (may affect search engine robots).
  - Download limits.
  - Geographic location of the server
    - Engaging by search engines,
    - Providing services from a US server to the PL market may result in incorrect results in the SERP.

Recognition of competition

- Companies that use illegal tricks:
  - usually appear and disappear in the SERP;
  - https://www.google.com/webmasters/tools/spamreport
  - In the long run this approach does not ensure success.
  - A genuine case of http://findgiftcards.com/ ("gift certificates"), which has cooperated with http://123counters.com/ providing free counters and (secretly) linking to the aforementioned site. Effect: Complete removal from Google.

Recognition of competition (2)

- Which companies for "our" keywords are higher in the SERP?
  - Are their sites completely indexed?
    - `site:`ichdomena.com
    - http://www.netconcepts.com/urlcheck/
  - Unique, relevant (keywords) page titles.
  - Do their pages have quite high *PageRanks*?
  - Do link descriptions contain keywords?

Recognition of competition (3)

- What keywords do they use?
  - Titles of pages,
  - Meta tag `keywords`
- Who links to their sites?
- What technologies do they use?
  - http://www.netcraft.com
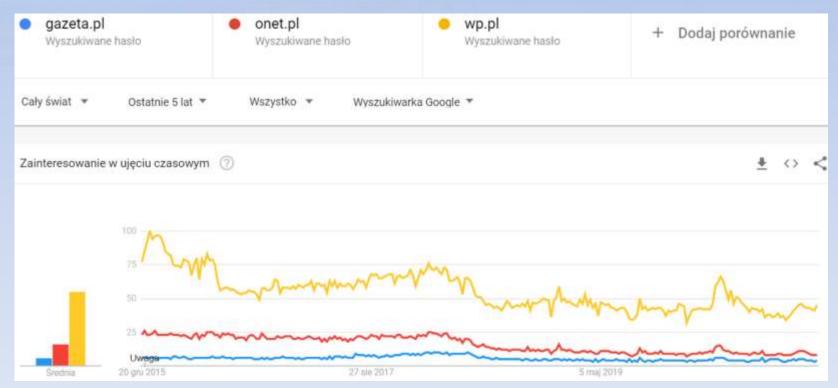- Do not use *cloaking* (different versions of site for different recipients) - check the search engine cache.

Recognition of competition (4)

- Evaluation of internet traffic of competing portals
  - https://www.similarweb.com/
  - http://www.quantcast.com/
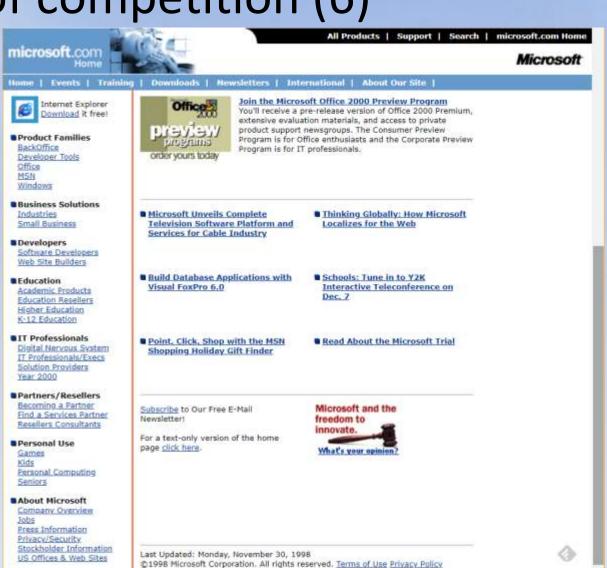 - http://www.alexa.com/

Recognition of competition (5)

- Evaluation of internet traffic of competing portals
  - o http://trends.google.com (2020-12)

Recognition of competition (6)

- How the competition site (and their SEO) changed over time?
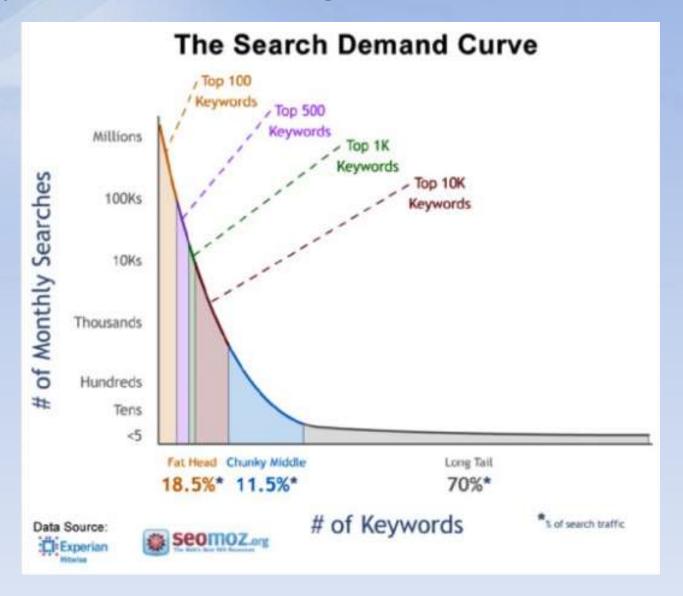 - http://www.archive.org/

Microsoft site

on 05-12-1998

Changes in the portal, and SEO

- New content, categories, functionality (services).
- Another domain.
- Changes to the URL structure of your site.
- Changing the CMS system.
- Collaboration with other entities for links.
- Changes to the navigation of the portal.
- Redirection.

Keywords

- Properly chosen keywords are one of the most important aspects of SEO.

- Should be a reflection of the words / phrases that users use during the search.

- It is difficult to optimize the portal for the most popular keywords.

- Therefore, it is worth using the terms of the so-called long tail (long tail).

Keywords from long tail

The Search Demand Curve

Keywords from long tail (2)

- The tail contains hundreds of millions of queries, only a small part of which repeats itself.
- As a result, it generates up to 70% of traffic.
- Therefore, it is worth trying to include such queries.

Keywords from long tail (3)

- Official Google information about "premiere" (first time seen) queries:
  - year 2007: 20 – 25% (source: [Udi Manber, Google's VP of Engineering](#)),
  - year 2013: 15% (source: [cnet](#)),
  - year 2017: 15% (source: [Google](#)),
- There is no easy way to find them. However, you can try to extract keywords from sites that are good at SERP:
  - Search the search engines for the most popular keywords,

Keywords from long tail (4)

- *Continued:*

  - For every page in the top 10 - 30, extract the text. You can also do a separate search, for example, discussion forums (Google parameter: `inurl:forum`);

  - We remove the expressions that are already in our resources;

  - Sort by most popular.

Create a keyword list

- You can start, for example, from brainstorming with your colleagues.
- We create a list of words and / or phrases (several dozens - several hundred items) containing up to 3 words.
- We then add synonyms (e.g. using a dictionary).
- We assign them to a certain tree hierarchy
- We add higher grade categories (for which our products / services are special cases).

Create a keyword list (2)

- Analyze the content of your existing portal and if possible add the words found there.
- Check the industry pages (e.g. newspapers) and add the expressions used there.
- Add the names of your products, e.g. models.
- Ask your team to come to our customer and try to formulate some queries - we add them to the list.

Create a keyword list (3)

- Ask for the same people not related to your company.
- Analyze the portals of your competitors:
  - What words / phrases do they use?
  - Look for similar information in published articles.
- Check social networking sites and blogs.

Create a keyword list - tools

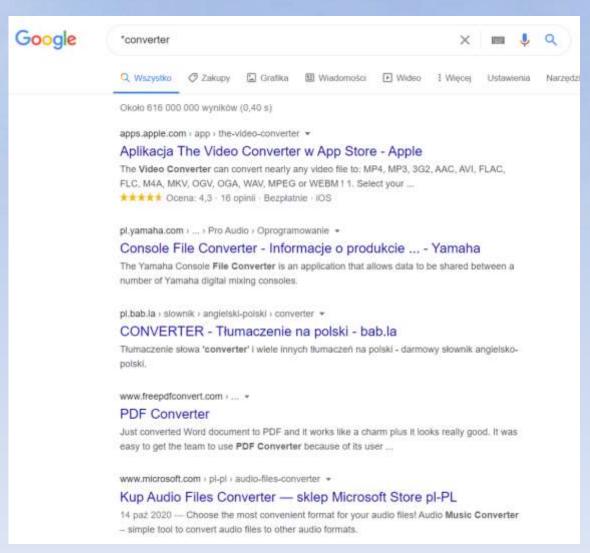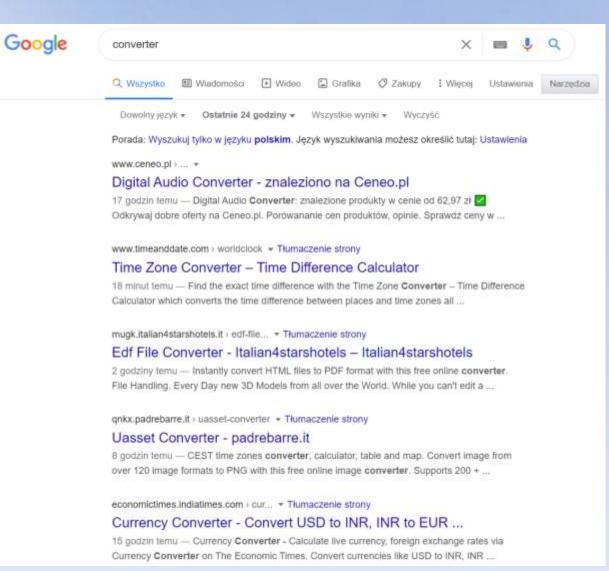● Related terms, synonyms (the cloud on the left) http://search.yippy.com

Create a keyword list – tools (2)

● Word associations:

  o Use the operator *

Create a keyword list – tools (3)

● Recent popularity of the term:

   o Use a date range operator (advanced search on Google)
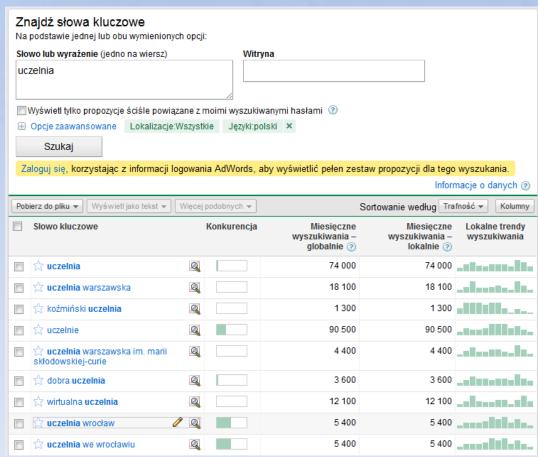
Create a keyword list – tools (4)
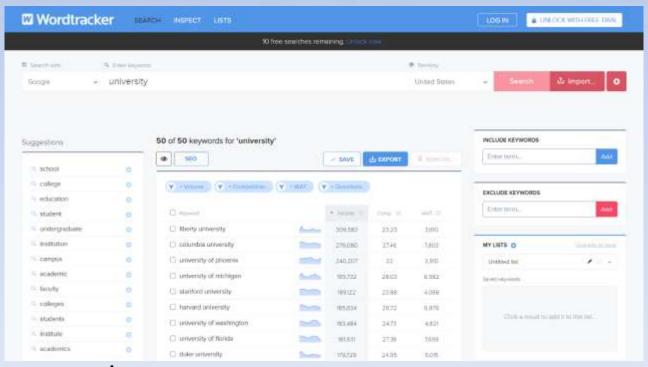
- https://adwords.google.com/select/KeywordToolExternal
  - o Word association,
  - o Statistics,
  - o Trends
  - o Costs (ads).

Create a keyword list – tools (5)

- http://www.wordtracker.com/
  - The most popular phrase containing the word,
  - approximate, daily number of searches containing the word;
  - Analysis and comparison of data including Google, Microsoft, Yahoo;
  - Phrases associated with the given word;

Create a keyword list – tools (6)

- http://www.wordtracker.com/ - *continued*
  - List of related keywords along with number of occurrences;
  - Reports: 1000 most popular expressions from last 48h;
  - The data comes from a few less popular search engines.

Create a keyword list – tools (7)

- https://www.spyfu.com/
  - a lot of different information,
  - also a free version.

Create a keyword list – tools (8)

- Other tools:
  - http://www.keyworddiscovery.com/
  - http://www.google.com/trends
  - https://monitorbacklinks.com/
  - https://www.semrush.com/

Create a keyword list – tools (9)

- Tools non-orginating from search engines are based on different data and / or estimations.
- Therefore, they are not always reliable in the context of the entire web.
- Hence, it is better to pay attention to certain trends instead of specific numbers.

Which keywords to choose?

- Unfortunately, there is no simple way that would be 100% effective.
- Value:
  - Usability in the context of our portal,
  - How much will we gain when we "aim" at this word?
- Relevance,
  - How much is the word related to the product / service?

Which keywords to choose? (2)

- Relevance – *continued*
  - Assuming that you click on the SERP result, what is the chance that it will perform the action we expect?
  - How many people will be disappointed with our content?
- Conversion rate (how many visitors will perform "advertised" action?).
  - It does not always have to be related to spending money - it could be, for example, creating a link to our site.

Which keywords to choose? (3)

● Conversion rate - continued

    ○ http://www.seomoz.org/keyword-difficulty/

    ○ Clicking in SERP depending on the place (AOL)

Position in SERP	„click" factor[%]
1	42,1
2	11,9
3	8,5
4	6,1
5	4,9

Source: http://www.webuildpages.com/jim/click-rate-for-top-10-search-results/

Which keywords to choose? (4)

- Conversion rate - continued
  - CR is usually about 50% better for ads than for "organic" results (source: Enquisite).
  - Conversion optimization involves performing different versions of the page and comparing the results (A / B tests).

Friendly portal for search engines

- First of all, the portal should be user friendly (high usability).

- Otherwise users will not use it and recommend it to others (that is, they will not link to our portal).

- Fortunately, there is no contradiction in the friendliness of:
  - users,
  - search engines.

Friendly portal for search engines (2)

- Let's make sure that search engines have access to the **entire** portal. The following solutions may cause problems:
  - Content hidden behind forms;
  - Links in JavaScript (sometimes working);
  - ~~Flash~~ links (e.g. fancy menus), ~~Java~~, etc.
  - Links on pages containing hundreds of links (probably not all will be processed).

Friendly portal for search engines (3)

- Let's make sure that search engines have access to the **entire** portal – *continued*:
  - Links in PowerPoint, PDF documents;
  - Links to pages blocked by:
    - *Robots.txt* file,
    - `NoFollow` **attribute**,
    - **Meta tag** `Robots`.
  - **Links in** `iframes` **and** `frames`.

Friendly portal for search engines (4)

- Make a *sitemap.xml* file.
  - Contains information (including links) about subpages on our site.
  - It is especially useful when a page can not be visited for the above reasons.
  - Detailed description: http://www.sitemaps.org/
  - Useful tools:
    - https://www.xml-sitemaps.com/ ,
    - plugins for popular CMSs.

Friendly portal for search engines (5)

- Proper structure of information (it is worth to use different prototypes – see the lecture on GUI Usability).
  - categories,
  - subcategories,
  - not much nesting (max. 3 - 4),
  - set up a maximum of 100 links per page,
  - avoid paging if possible or use the `<nav>` tag.

Friendly portal for search engines (6)

- Note on subdomains: should cover different areas.

- Friendly URLs
  - include business information,
  - use "-" instead of ". "
  - avoid parameters, session id, etc.

- Note on AJAX and JavaScript. Let's make sure that each page in the portal is also available in a classic way (*Unobtrusive Javascript*).

Friendly portal for search engines (7)

- Avoid frames (`frame`, `iframe`).
  - o Search engines usually treat them as outside content.
  - o Therefore, they usually do not index it.
- Adding content from session IDs and cookies may result in inaccessibility of pages for search engines.
- Unique URL for each page.

To be continued…