



POLISH-JAPANESE ACADEMY  
OF INFORMATION TECHNOLOGY

# Content Management Systems (CMS)

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

Mariusz Trzaska, Ph. D.

[mtrzaska@mtrzaska.com](mailto: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/](http://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)

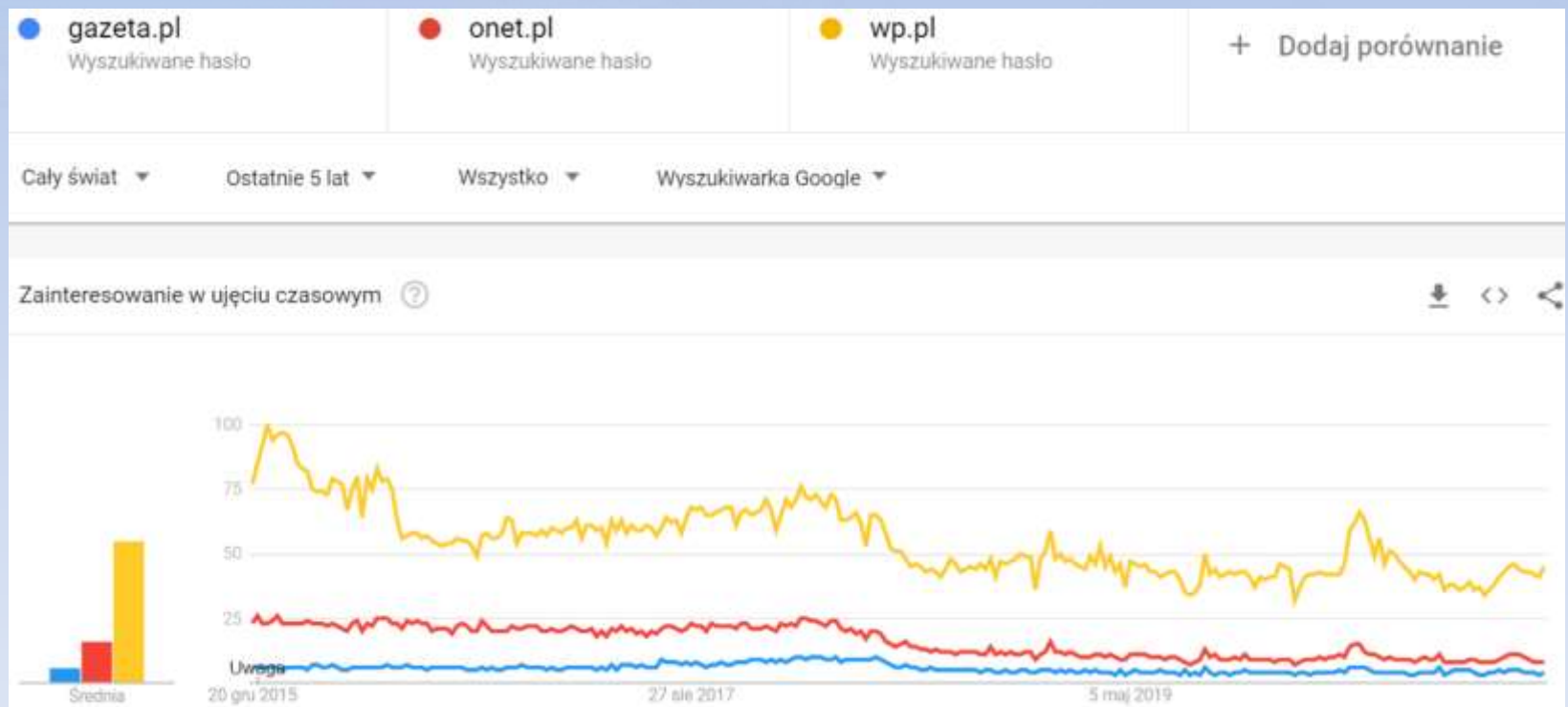
- Evaluation of internet traffic of competing portals
  - <https://www.similarweb.com/>
  - <http://www.quantcast.com/>
  - <http://www.alexa.com/>



# Recognition of competition (5)

- Evaluation of internet traffic of competing portals

- <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

The screenshot shows the Microsoft.com homepage as of May 12, 1998. The layout includes a top navigation bar with links for 'All Products', 'Support', 'Search', and 'microsoft.com Home'. Below this is a secondary navigation bar with links for 'Home', 'Events', 'Training', 'Downloads', 'Newsletters', 'International', and 'About Our Site'. A left sidebar contains a list of product families and user groups, including 'Product Families' (BackOffice, Developer Tools, Office, MSN, Windows), 'Business Solutions' (Industries, Small Business), 'Developers' (Software Developers, Web Site Builders), 'Education' (Academic Products, Education Resellers, Higher Education, K-12 Education), 'IT Professionals' (Digital Nervous System, IT Professionals/Execs, Solution Providers, Year 2000), 'Partners/Resellers' (Becoming a Partner, Find a Services Partner, Resellers Consultants), 'Personal Use' (Games, Kids, Personal Computing, Seniors), and 'About Microsoft' (Company Overview, Jobs, Press Information, Privacy/Security, Stockholder Information, US Offices & Web Sites). The main content area features a large advertisement for 'Office 2000 preview programs' with the text 'order yours today'. Below this are several news snippets, including 'Microsoft Unveils Complete Television Software Platform and Services for Cable Industry', 'Thinking Globally: How Microsoft Localizes for the Web', 'Build Database Applications with Visual FoxPro 6.0', 'Schools: Tune in to Y2K Interactive Teleconference on Dec. 2', 'Point, Click, Shop with the MSN Shopping Holiday Gift Finder', and 'Read About the Microsoft Trial'. At the bottom, there is a section for 'Subscribe to Our Free E-Mail Newsletter!' and a link to 'Microsoft and the freedom to innovate. What's your opinion?'. The footer contains the text 'Last Updated: Monday, November 30, 1998' and '©1998 Microsoft Corporation. All rights reserved. Terms of Use Privacy Policy'.

# Changes in the portal, and SEO

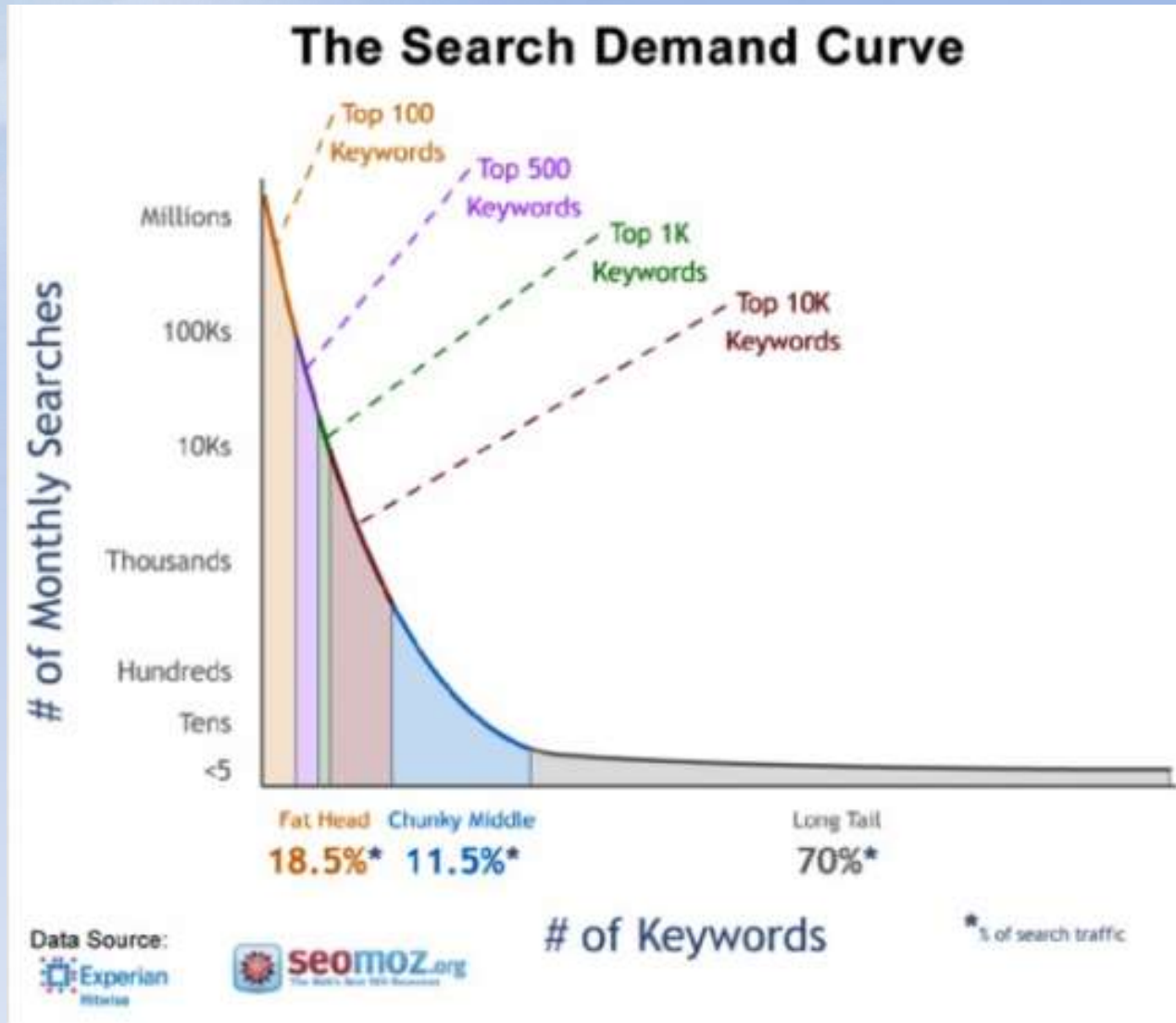
- New content, categories, functionality (services).
- Another domain.
- Changes to the URL structure of your site.
- Changing the CMS system.
- Collaboration with other entities for links.
- Changes to the navigation of the portal.
- Redirection.

# 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



# 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>

The screenshot shows the Yippy search engine interface. The search term 'video converter' is entered in the search bar. The results page displays a list of search results, including links to various video conversion tools and services. The left sidebar shows a navigation menu with categories like 'Sources', 'Sites', 'Time', and 'Topics'. The main content area lists search results with titles, snippets, and links to preview or view the full page.

**yippy** Web News Images Video

video converter Search

Results 1-20 of about 2,132 | Details

Sources Sites Time Topics

**Top 363 Clusters** remix

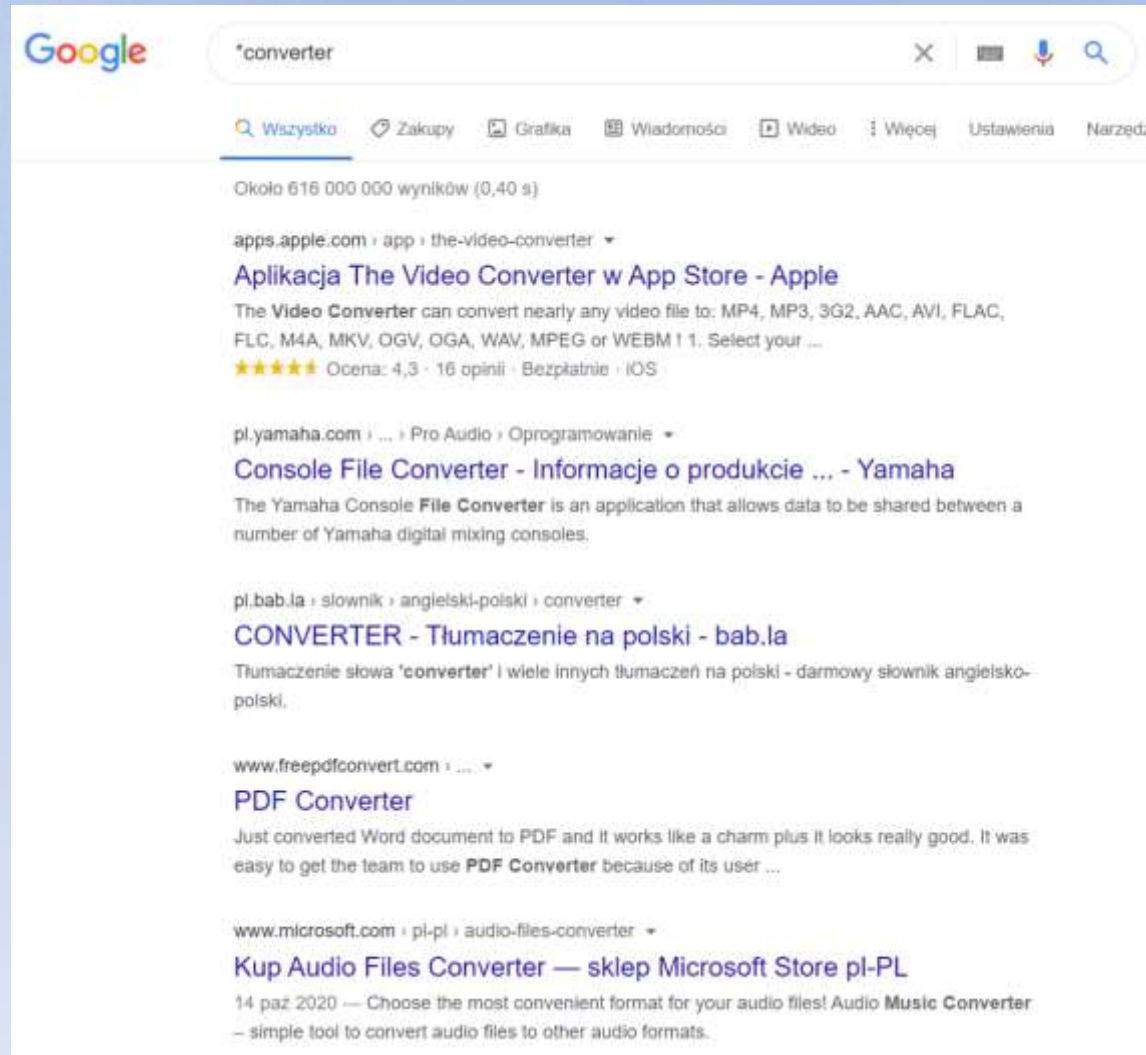
- + Download, Software (109)
- + iPod , DVD (56)
- + 126 下载, URex Video Converter 绿色特别版 2.0 (58)
- + Image (30)
- + Data Recovery (14)
- + Xilisoft Video Converter Ultimate (13)
- Scan, Manufacturer (6)
- + Players, Converters (8)
- + Adapters, Accessories (10)
- + Free Audio (8)
- + Broadcast, Audio (5)
- Features (5)
- + Extenders, VGA (6)
- + (7) دالووم, نرم افزار
- 七喜下载站, 官方软件下载, 绿色软件, 游戏下载, 最新 (7)
- Digital Video (4)
- + アプリやWebの疑問に答えるメディア (6)
- Format converters (4)
- Software Reviews (3)
- Optimized, Powerful Video Converter (2)
- Wav, MP3 , Wma (4)
- Tuner, TV (2)
- 0.1 (3)
- Signal processing (3)
- Matrix Switch (2)
- Avi To Mp3 (1)

1. [In Search Of Christian Entrepreneurship Home School Content - Kevin Cullis](#) [new window](#) [preview](#)  
... better rate of salvations than what he called "converters" (which focus on evangelization alone). The results were ...  
<https://finance.townhall.com/...ome-school-content-n2549546?amp=true> - [cache](#) - Yippy News
2. [Audio Tools Factory](#) [new window](#) [preview](#)  
MP3 Cutter Joiner, CD To MP3 WAV Maker, WAV MP3 Converter, MP3 RM Converter, Video To Audio converter, and Audio Tag Editor.  
[www.audiotoolsfactory.com](http://www.audiotoolsfactory.com) - [cache](#) - Yippy Index I
3. [Easy video conversion to any device - Free Video Converter](#) [new window](#) [preview](#)  
... and 64-bit versions. FINALLY A SIMPLE, FREE VIDEO CONVERTER Converting your videos to play optimally on a ... be complicated. Free File Converter is a powerful video converter without a lot of advanced configuration options you ...  
<https://www.freefileconverter.com> - [cache](#) - Yippy Index
4. [Super DVD Creator, Convert & burn AVI MP4 WMV DivX OGG to dvd with dvd menu and subtitle, Video Converter, rip DVD to ...](#) [new window](#) [preview](#)  
DVD Creator DVD Ripper Video Converter iPhone iPod PSP Home Products Download Buy now Guides Contact us Super DVD Creator the easy way to burn your ... 2003, Super DVD Creator is not only a video converter and DVD burner, it is a professional all- ... Buy Now Apple Video Tools, DVD ripper and video converter for your iPod iPhone and Apple-TV . [ iPhone/ ...  
[alldj.com](http://alldj.com) - [cache](#) - Yippy Index
5. [ePVision](#) [new window](#) [preview](#)  
Offers a variety of HDTV tuner, TVs, cameras, and video converters or enhancer.  
[www.epvision.com](http://www.epvision.com) - [cache](#) - Yippy Index I
6. [AnySoft - Video Converter, Video Converter, Audio Converter, Photo DVD Maker, PDFMate Converter, Photo Slideshow Creator](#) [new window](#) [preview](#)  
Ultimate An all-in-one Video Recorder, DVD Converter, Video Converter and Audio Extractor



# Create a keyword list – tools (2)

- Word associations:
  - Use the operator \*



The screenshot shows a Google search for the keyword "converter". The search bar at the top contains the text "\*converter". Below the search bar, there are navigation tabs for "Wszystko", "Zakupy", "Grafika", "Wiadomości", "Wideo", "Więcej", "Ustawienia", and "Narzędzia". The search results indicate approximately 616,000,000 results found in 0.40 seconds.

The first result is from [apps.apple.com](https://apps.apple.com), titled "Aplikacja The Video Converter w App Store - Apple". The description states: "The Video Converter can convert nearly any video file to: MP4, MP3, 3G2, AAC, AVI, FLAC, FLC, M4A, MKV, OGV, OGA, WAV, MPEG 1 1. Select your ...". It has a rating of 4.3 stars based on 16 reviews and is available for free on the iOS App Store.

The second result is from [pl.yamaha.com](https://pl.yamaha.com), titled "Console File Converter - Informacje o produkcie ... - Yamaha". The description states: "The Yamaha Console File Converter is an application that allows data to be shared between a number of Yamaha digital mixing consoles."

The third result is from [pl.bab.la](https://pl.bab.la), titled "CONVERTER - Tłumaczenie na polski - bab.la". The description states: "Tłumaczenie słowa 'converter' i wiele innych tłumaczeń na polski - darmowy słownik angielsko-polski."

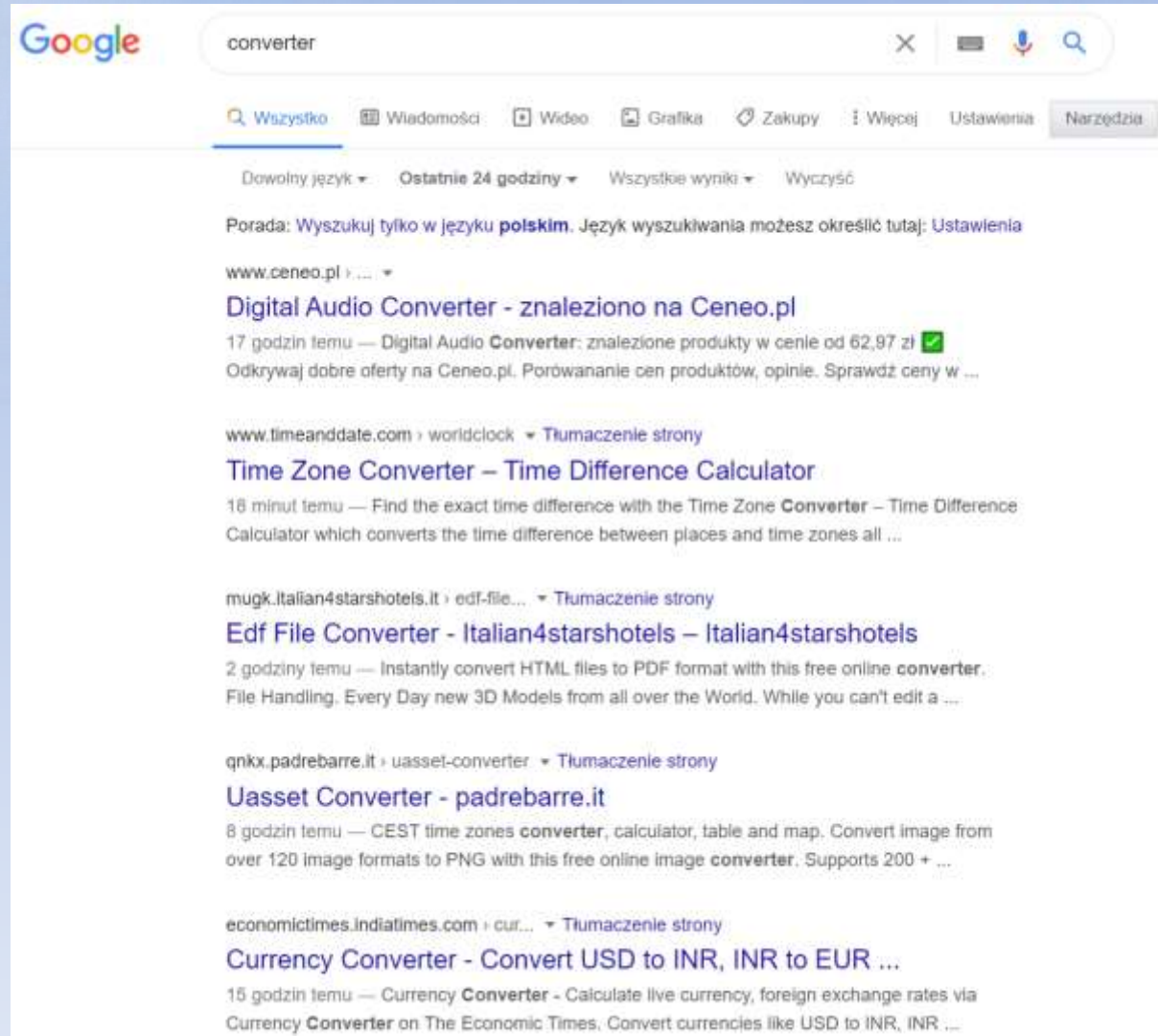
The fourth result is from [www.freepdfconvert.com](https://www.freepdfconvert.com), titled "PDF Converter". The description states: "Just converted Word document to PDF and it works like a charm plus it looks really good. It was easy to get the team to use PDF Converter because of its user ..."

The fifth result is from [www.microsoft.com](https://www.microsoft.com), titled "Kup Audio Files Converter — sklep Microsoft Store pl-PL". The description states: "14 paź 2020 — Choose the most convenient format for your audio files! Audio Music Converter – simple tool to convert audio files to other audio formats."

# Create a keyword list – tools (3)

## ● Recent popularity of the term:

- Use a date range operator (advanced search on Google)



The screenshot shows a Google search for the keyword "converter". The search results are displayed in Polish. The first result is "Digital Audio Converter - znaleziono na Ceneo.pl" with a date range filter of "17 godzin temu". The second result is "Time Zone Converter – Time Difference Calculator" with a date range filter of "18 minut temu". The third result is "Edf File Converter - Italian4starshotels – Italian4starshotels" with a date range filter of "2 godziny temu". The fourth result is "Uasset Converter - padrebarre.it" with a date range filter of "8 godzin temu". The fifth result is "Currency Converter - Convert USD to INR, INR to EUR ..." with a date range filter of "15 godzin temu".

# Create a keyword list – tools (4)

● <https://adwords.google.com/select/KeywordToolExternal>

- Word association,
- Statistics,
- Trends
- Costs (ads).

Znajdź słowa kluczowe  
Na podstawie jednej lub obu wymienionych opcji:

Słowo lub wyrażenie (jedno na wiersz)  Witryna

Wyświetl tylko propozycje ściśle powiązane z moimi wyszukiwanymi hasłami [?](#)

[+](#) Opcje zaawansowane Lokalizacja: Wszystkie Języki: polski [×](#)

Zaloguj się, korzystając z informacji logowania AdWords, aby wyświetlić pełen zestaw propozycji dla tego wyszukiwania. [Informacje o danych](#) [?](#)

Pobierz do pliku  Więcej podobnych

<input type="checkbox"/>	Słowo kluczowe	Konkurencja	Miesięczne wyszukiwania – globalnie <a href="#">?</a>	Miesięczne wyszukiwania – lokalnie <a href="#">?</a>	Lokalne trendy wyszukiwania
<input type="checkbox"/>	☆ uczelnia	<input type="text"/>	74 000	74 000	
<input type="checkbox"/>	☆ uczelnia warszawska	<input type="text"/>	18 100	18 100	
<input type="checkbox"/>	☆ koźmiński uczelnia	<input type="text"/>	1 300	1 300	
<input type="checkbox"/>	☆ uczelnie	<input type="text"/>	90 500	90 500	
<input type="checkbox"/>	☆ uczelnia warszawska im. marii skłodowskiej-curie	<input type="text"/>	4 400	4 400	
<input type="checkbox"/>	☆ dobra uczelnia	<input type="text"/>	3 600	3 600	
<input type="checkbox"/>	☆ wirtualna uczelnia	<input type="text"/>	12 100	12 100	
<input type="checkbox"/>	☆ uczelnia wrocław	<input type="text"/>	5 400	5 400	
<input type="checkbox"/>	☆ uczelnia we wrocławiu	<input type="text"/>	5 400	5 400	



# 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;

The screenshot displays the Wordtracker website interface. At the top, there is a navigation bar with 'Wordtracker', 'SEARCH', 'INSPECT', and 'LISTS' tabs. A 'LOG IN' button and a 'UNLOCK WITH FREE TRIAL' button are also visible. Below the navigation bar, a search bar contains the text 'Google' and 'university', with 'United States' selected as the location. A 'Search' button and an 'Import...' button are present. The main content area is titled '50 of 50 keywords for 'university'' and features a table of keywords with associated data. A 'Suggestions' sidebar on the left lists related terms like 'school', 'college', and 'education'. On the right, there are sections for 'INCLUDE KEYWORDS' and 'EXCLUDE KEYWORDS', each with an 'Add' button. At the bottom right, there is a 'MY LISTS' section with an 'Untitled list' and a 'Save keywords' button.

Keyword	Volume	Competition	Cost Per Click	Quality Score
liberty university	306,583	23.23	3.00	
columbia university	276,080	27.46	7.00	
university of phoenix	240,307	22	2.80	
university of michigan	193,722	28.03	8.92	
stanford university	189,122	23.98	4.09	
harvard university	185,034	26.72	8.85	
university of washington	183,484	24.73	4.82	
university of florida	181,631	27.39	7.99	
duke university	179,729	24.95	5.00	

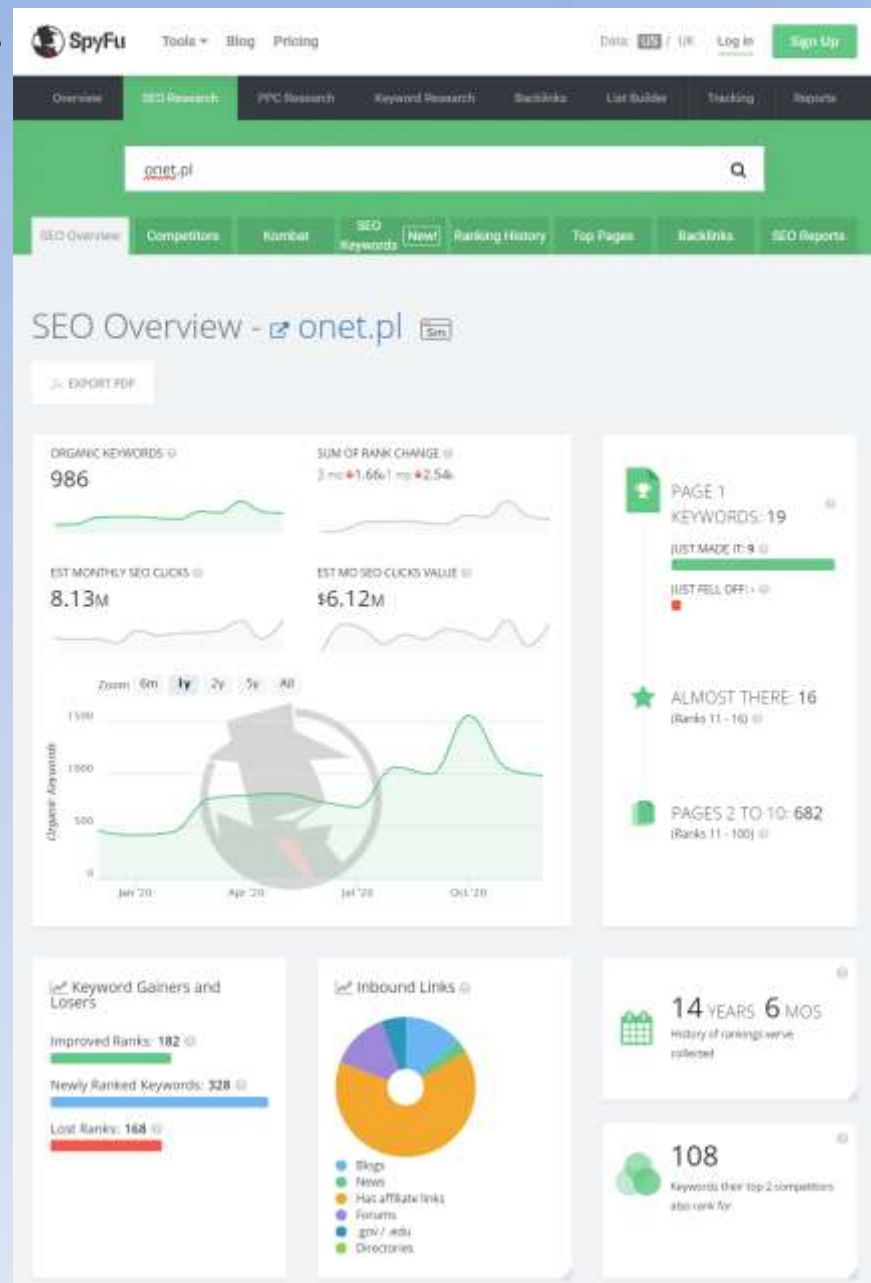
# Create a keyword list – tools (6)

- <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/>

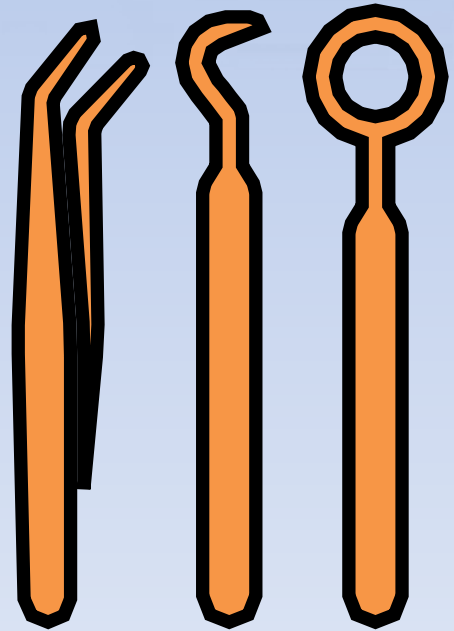
- a lot of different information,
- also a free version.



# 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-originating 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`).
  - Search engines usually treat them as outside content.
  - 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...*