



POLISH-JAPANESE ACADEMY  
OF INFORMATION TECHNOLOGY

# Content Management Systems (CMS)

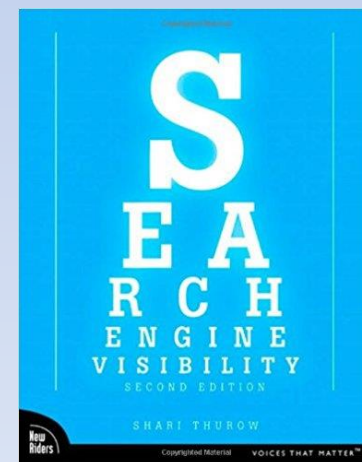
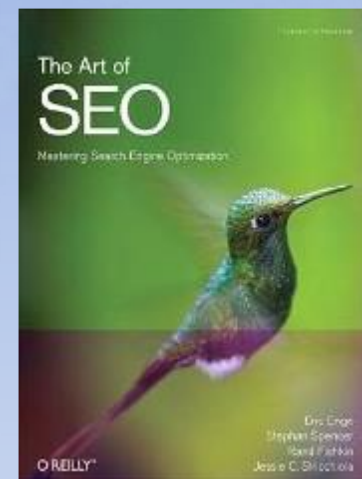
## Lecture 11: Search engines and content management

Mariusz Trzaska, Ph. D.

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

# References

- Eric Enge, Stephan Spencer, Rand Fishkin, Jessie C Stricchiola: The Art of SEO: Mastering Search Engine Optimization (Theory in Practice). O'Reilly Media.
- Shari Thurow: Search Engine Visibility (2nd Edition). New Riders Pub.



# Agenda

- What is SEO?
- Search engines and their queries
- Factors determining SEO success
- Strategic SEO goals
- Audit of an existing portal
- Recognition of competition
- Keywords
- Links
- Summary

# What is SEO?

- SEO (*Search Engine Optimization*) – **process** which purpose is to make the published content easily discoverable by search engines.
- Positioning is designed to place a particular website high in the search results.
- Both concepts are linked together.

# What is SEO? (2)

- There are two approaches that can be used:
  - „Dark side of the power” (*black hat*)
    - It’s purpose is to „cheat” the search engine
    - Usually provides short-term effects.
  - „Bright side of the power” (*white hat*)



# What is SEO? (3)

- Contrary to popular belief, there are no effective tricks to ensure a quick and lasting success.
- Properly implemented SEO is related to:
  - Continuous delivery of fresh content to the portal,
  - Properly formed, technically correct pages,
  - Marketing,
  - Use of appropriate technology.
- So with (unceasing) hard work.

# Modern SEO

- Includes various aspects:
  - technical,
  - marketing
  - business.
- Clear (measurable) targets should be set.
- This is a **process** that never really ends.
- It is quite long, so it requires patience.
- It is usually based on assumptions, because search engine developers are hiding/guarding their secrets.

# Search engine market

Google	69,80%
Microsoft (Bing)	13,31%
Baidu	12,53%
Yahoo	2,11%
Yandex	1,19%

Global market 2020-12. Source: <http://www.netmarketshare.com/>

Official Google materials „*Creating websites that work well with Google*“:  
<https://developers.google.com/search/docs/beginner/get-started>



# Typical query

Number of words in the query	% queries
1	25,32
2	24,96
3	19,80
4	13,17
5	7,53
6	4,04
7	2,15
8	1,19

Source: comScore (2009-03)

# Search goals

- To provide the best results, you must specify a search target.
- It depends on the type of information that should be returned.
- Potential goals:
  - Navigation,
  - Informative,
  - Transactional.

# Search goal - navigation

- A user would like to visit a particular site without knowing the exact address, such as the site of a particular hotel.
- Possibilities:
  - Pull the user from the original destination.
  - Obtain an auxiliary traffic.
- Potential value:
  - Quite low,
  - With the ability to change the search destination.

# Search goal - informative

- Very broad category, for example information about:
  - weather,
  - the latest events,
  - products,
  - services,
  - scientific issues.
- User does not care about a particular site.

# Search goal – informative (2)

## ● Possibilities:

- Attracting visitors,
- Creating links to our site
- Interest of different customer groups (depending on the topic).

## ● Potential value:

- Average,
- These types of queries, connected to e.g. products may have a high value.

# Search goal - transactional

- Performing some activities (not necessarily commercial ones), e.g.
  - Create a mail account,
  - Registering in the discussion forum,
  - Making a transfer,
  - Book a table in a restaurant,
  - Shopping,
  - Subscribe to the RSS feed,
  - Job requests.

# Search goal – transactional (2)

- Possibilities:

- Quite big

- Potential value

- Quite big

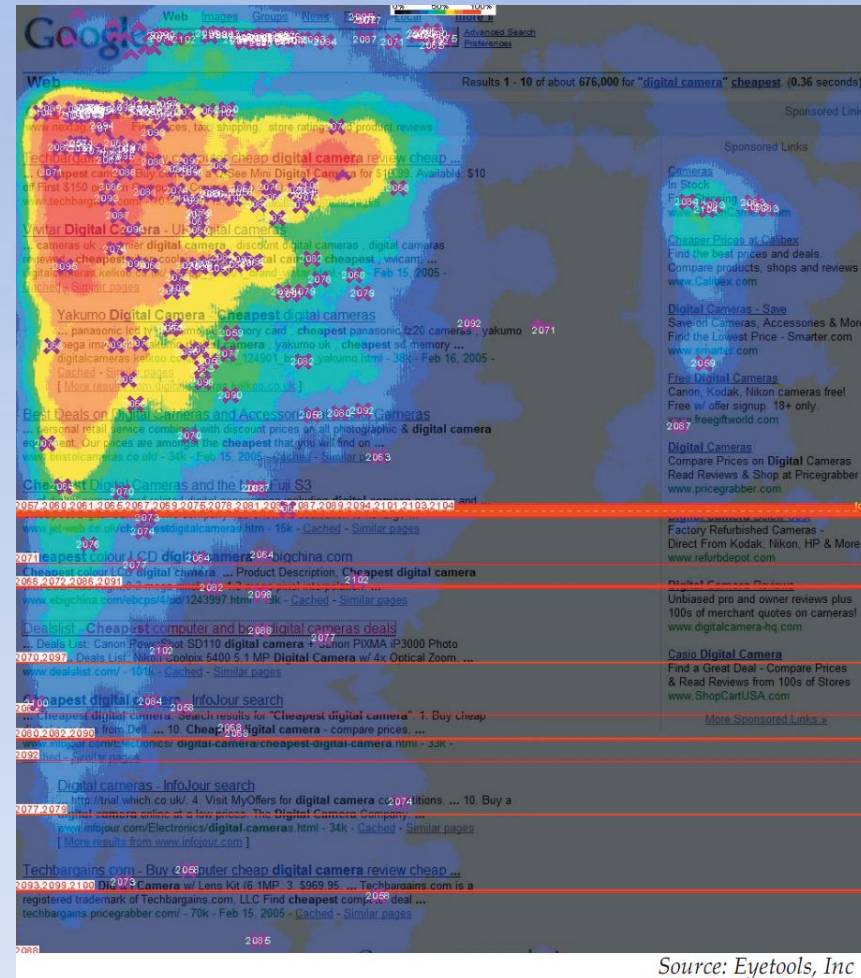
# How do people search?

- They usually work iteratively.
- They use different sources of information, intensively navigating.
- 30% of online transactions are completed more than 24 hours after the initial search.



# Evaluation of search results

- Results analysis – *eyetracking*.
  - „F” shape.
- iProspect and Jupiter Research:
  - 62% of users choose results from 1st page.
  - 90% of the first 3 pages.



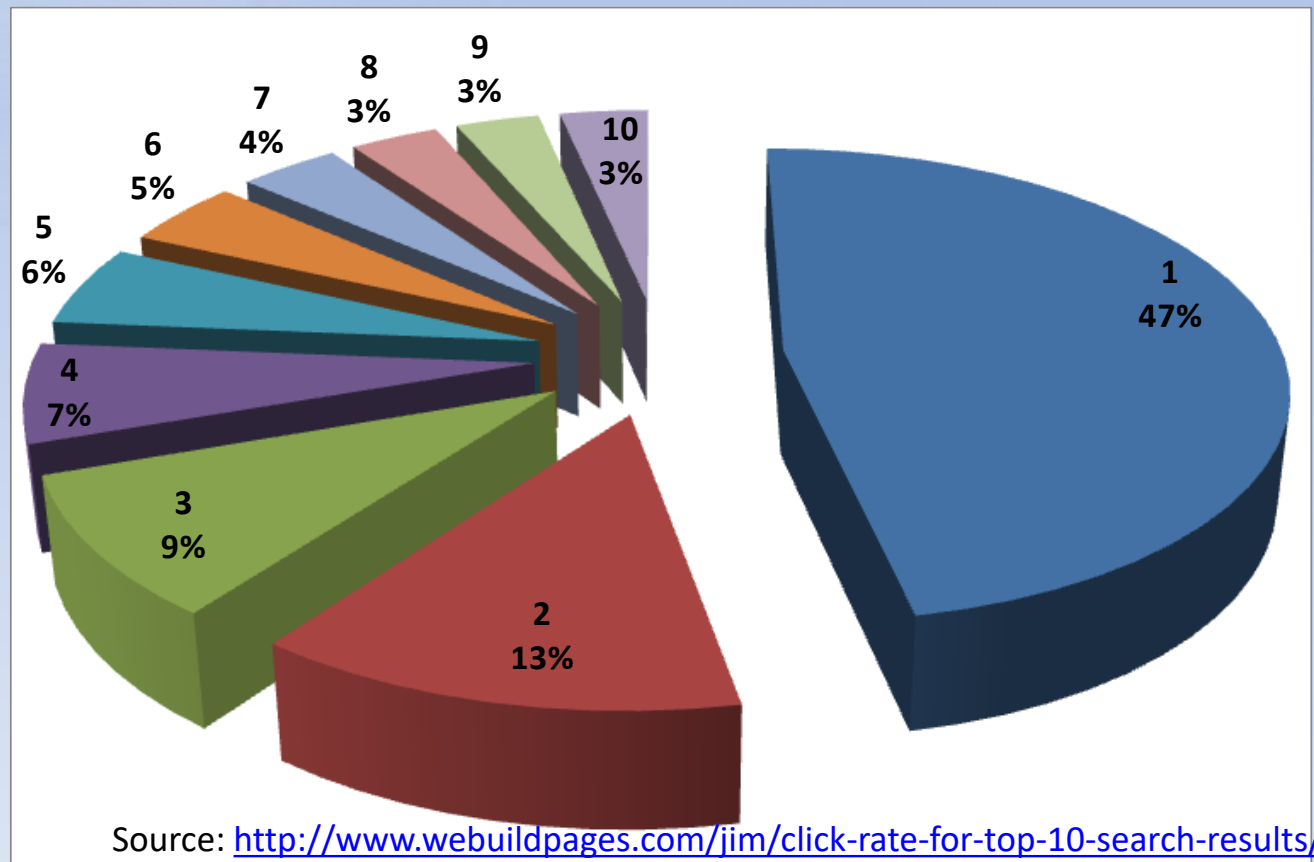
Source: Eyetools, Inc

# Evaluation of search results (2)

- iProspect and Jupiter Research from –  
*continued:*
  - 41% of users change criteria when they do not find answers on the first page,
  - 88% do the same after 3 pages,
  - 36% agree that finding a company on the 1st page ensures its high position in the area of activity.

# Evaluation of search results (3)

- Study conducted on AOL data ("clickability" depending on position in results):



# Evaluation of search results(4)

Position	Visibility - „natural” results[%]	Visibility - „sponsored” results (ads) [%]
1	100	50
2	100	40
3	100	30
4	85	20
5	60	10
6	50	10
7	50	10
8	30	10
9	30	

Source: Enquiro, Didit, Eyetools

# Evaluation of search results (5)

- 85% of users click on "natural" results.
- The first 4 places of sponsored results correspond to 7-10 sites in natural results (visibility and clickability)
- Studies show that the place in natural results (# 3 or better) is 2-3 times more valuable than the sponsored ones.

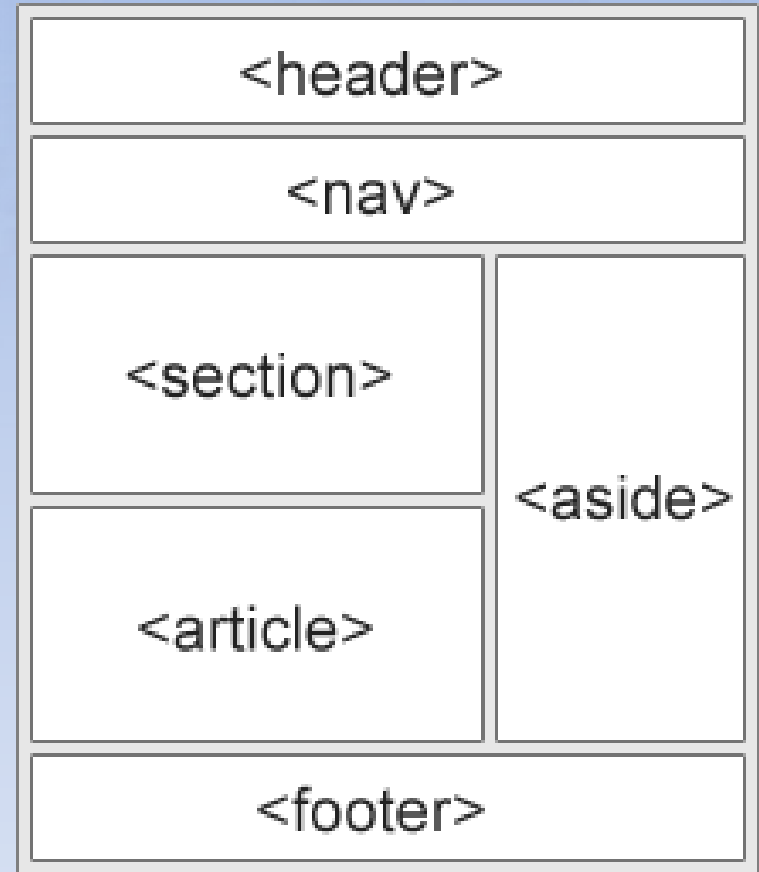
Source: Enquiro, Didit, Eyetools

# Visibility of page elements for the search engine

- Page title,
- Meta tags,
  - Keywords
  - Description
- Alt attribute for images,
- Noscript tag,

# Visibility of page elements for the search engine (2)

- The content of the page. HTML5 semantic (Structural) elements:
  - `<article>`
  - `<aside>`
  - `<details>`
  - `<figcaption>`
  - `<figure>`
  - `<footer>`
  - `<header>`
  - `<main>`
  - `<mark>`
  - `<nav>`
  - `<section>`
  - `<summary>`
  - `<time>`



[https://www.w3schools.com/html/html5\\_semantic\\_elements.asp](https://www.w3schools.com/html/html5_semantic_elements.asp)



# Search engines and semantic relationships

- Modern search engines, to a certain extent, recognize semantic relationships between words.
- For this purpose, special dictionaries (thesaurus) and fuzzy set theory (fuzzy set) are used.
- Types of searches using semantic relationships:
  - Proximity - documents containing a specific phrase.
  - Fuzzy logic - used taking into account syntactic errors.



# Search engines and semantic relationships (2)

## ● Types of searches – *continued*:

- Boolean - use of operators AND, OR, NOT.
- Term weighting - varying the weight of particular words in a query (e.g. skipping very popular expressions).

# Search engines and semantic relationships (3)

## ● *Latent Semantic Analysis (LSA)*

- In order to improve search results, we use huge databases of indexed words.
- This allows you to determine which words are associated with each other.
- On this basis, their dependencies can be inferred.
- So far, these are relatively simple information, for example, the search engine is able to say that some words are close to each other but does not "understand" the nature of the relationship.

# Search engines and semantic relationships (4)

- *Latent Semantic Indexing (LSI)*

- Using LSI, the search engine is able to identify relationships between sites (containing related words).
- Also used in contextual advertising (such as Google AdSense).

- Requests to include when creating content

- Using synonyms

# Link analysis

- The number of links leading to a particular site and their source are taken into account.
- Sites-authorities.
  - Specifies a collection of sites related to a topic;
  - Among them is the site that receives the most links from sites in this collection.
- A site with lots of links but low quality pages is also rated low.

## Link analysis (2)

- *Link neighborhood*. A group of sites related to similar topics.
- The group in which a particular site is located also projects on its value.
- The *anchor* text is quite important in evaluating the content of the site.
- It is also important that the link is located on a similar page as the landing page.

# Interpretation of the query

- Sometimes search engines have trouble identifying a search target.
- In this situation, results from different categories are included in the results.
- As a result, they may find sites with lower "value", but differentiating the categories received.
- Sometimes there is a hint that makes it possible to refine the results.

# Factors determining SEO success

- SEOmoz research

(<http://www.seomoz.org/article/search-ranking-factors>) is among the top SEO experts.

- Positive factors

- Well written content describing the link (*anchor*) leading to our site: 73%,
- Link popularity to our website: 71%,
- Diversity of links: 67%,
- Keywords in page title: 66%.

# Factors determining SEO success (2)

- Positive factors – *continued*

- Domain trust: 66%

- Negative factors:

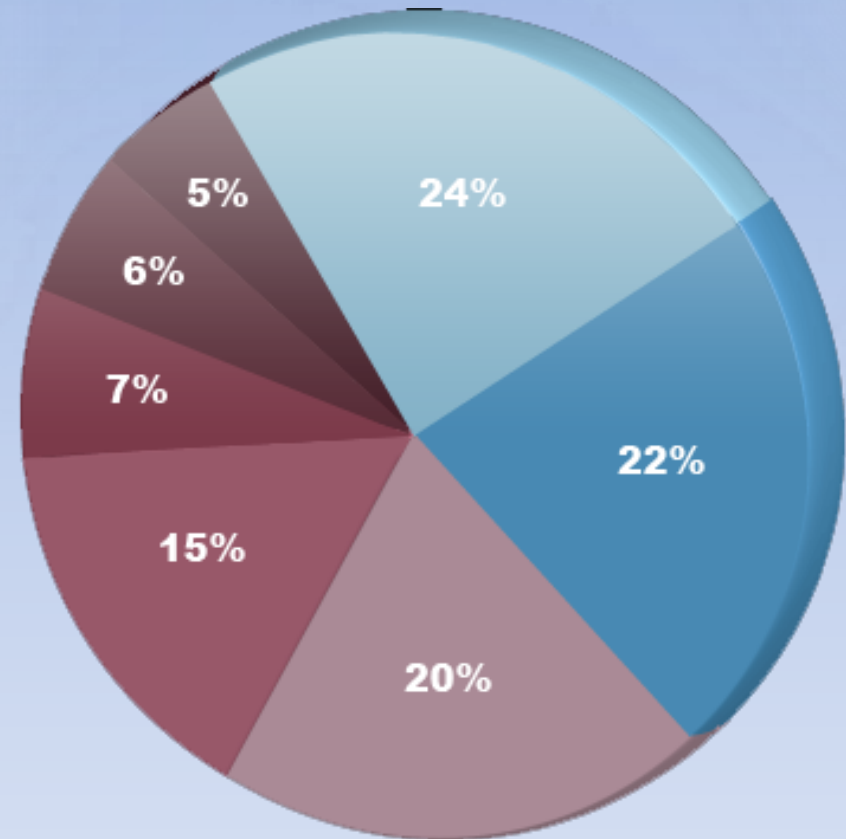
- Detected manipulation / cloaking attempts: 68%;
- Buying links from well known Brokers: 56%;
- Links to spam sites: 51%;
- Cheating with User Agent: 51%;
- Frequent server inaccessibility: 51%.



# Factors determining SEO success (3)

## ● Ranking algorithm:

- 24%: Domain trust,
- 22%: Link popularity,
- 20%: Text describing the link (*anchor*)
- 15%: Use keywords on the page,
- 7%: Generated traffic,
- 6%: Social Metrics,
- 5%: Domain registration data,



Source:  
<http://www.seomoz.org/article/search-ranking-factors>

# Strategic SEO goals

## ● Visibility of the brand

- The results at the top of the SERP are received by consumers as a quality mark.
- Concrete business names usually do not require any special action from the SEO side (unless they are popular).
- The case is complicated when we want to associate product types (e.g. "laptop") with our brand.

# Strategic SEO goals(2)

- Increase traffic to your site
  - Encourage existing customers
  - Acquiring new ones
  - Include different queries entered by users.
- ROI (Return On Investment)
  - Acquire "proper" traffic
  - Usually SEO provides better value for money than other media (press, radio, TV).

# Strategic SEO goals (3)

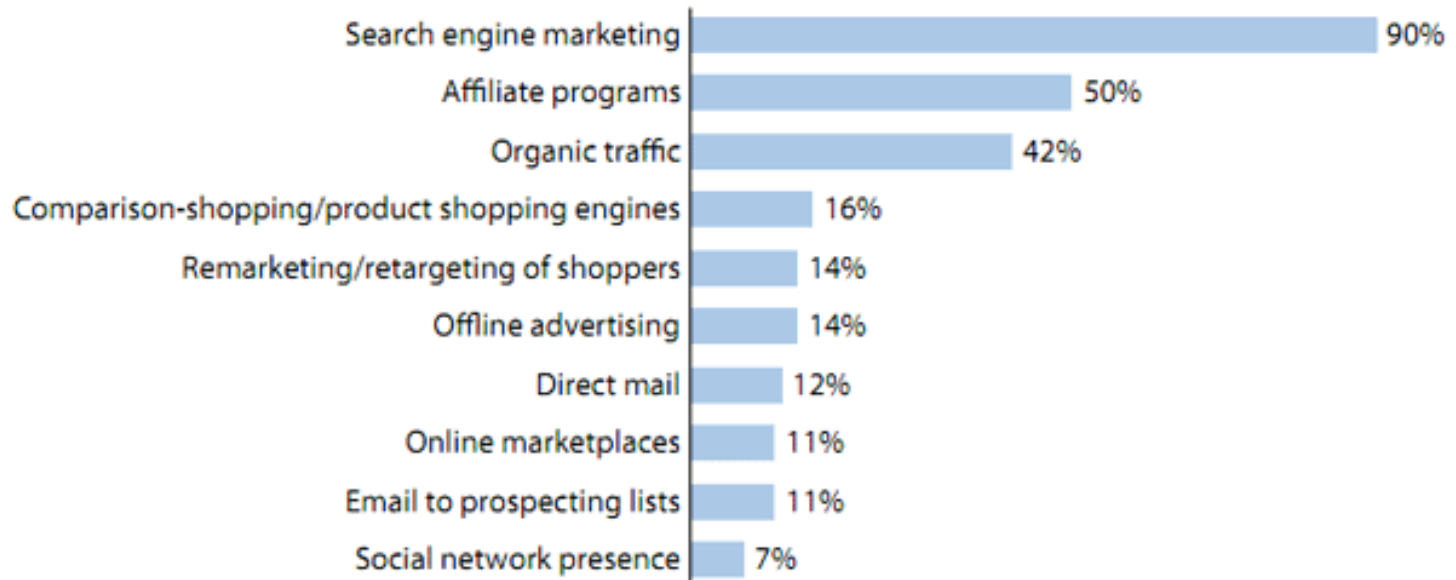
- ROI (Return On Investment) – *continued*.
  - Research conducted by SEMPO (*What are the most effective forms of advertising*):
    - 63% - *product placement*,
    - **50% - Natural / organic SEO,**
    - 44% - email marketing (not spam)
    - 13% - conferences and exhibitions,
    - 12% - *Public Relations*,
    - 11% - contextual text ads.

# Strategic SEO goals (4)

- Three most effective ways to win customers

## 1-2 Social networks rank last among 10 customer acquisition tactics

“Which of the following were your top three most effective sources used to acquire customers in 2009?”



Base: 102 online retailers

Source: “The State Of Retailing Online 2010,” a Shop.org study conducted by Forrester Research

58603

Source: Forrester Research, Inc.

# Perfect SEO strategy

- Unfortunately, there is no such thing.
- Every SEO plan must be tailored to the specific situation and include, among others:
  - Subject of promotion,
  - Target market,
  - Brand,
  - The structure and current content of the site.
  - Ease of modification of the above elements,
  - Resources for creating new content,
  - Competition.

# Creating SEO plan

- The incorporation of SEO should be made at the earliest possible stage of the portal design.
- Selected implementation technologies (or ready-made CMS) can affect SEO.
- Business aspects
  - Method of earning (revenue model),
  - Targeted customers.

# Creating SEO plan (2)

## ● Business aspects – *continued*

- The strategy used by the competition,
- Brand awareness (*branding*),
- Budget and schedule for content creation (not system),
- How to find information by potential customers.



# Own way for success

- Finding your own niche.
- Understanding
  - Who?
  - How?
  - What is he/she looking for?
- *Content Is King*
  - Without good content we have no chance for long-term success!
  - Content is a way to attract visitors (and gain links).
  - Thanks to that we can offer them our products / services.

# Own way for success(2)

## ● *Content Is King – continued*

- We may offer some interesting content for **our** customers, for example:
  - FAQ,
  - „How to”,
  - Video,
  - Galleries,
  - Services.
- To become a "leader" we have to offer something new!

# Own way for success (3)

- Be sure to know your users:
  - Men and women may use different keywords for the same thing;
  - Geographic diversification;
  - Education,
  - Earnings,
  - Culture,
  - ...

# The right team

- Leader
- Technical group
- Creative group
- Representatives (as advisors):
  - marketing,
  - ads
  - PR.

# Analysis of portal architecture information

## ● Technical aspects

- Dynamic URLs (without) with as few parameters as possible
- URLs without session ID and / or user ID
- Links or content used. Be careful with:
  - JavaScript,
  - ~~Java,~~
  - ~~Flash~~
- Content available after completing the form.

# Analysis of portal architecture information (2)

- Technical aspects – *continued*

- Temporary (302) redirects (usually better to use permanent one: 301).

- Structural aspects:

- Keywords,
- Cross-linking (inside the portal),
- Proper use of anchor text:
  - Relevant keywords,
  - Avoid "*More*", "*Click here*", etc.

# Analysis of portal architecture information (3)

## ● Structural aspects – *continued*.

### ○ Breadcrumbs:

- Usually, it naturally includes links with multiple keywords (e.g. categories).
- Usability improvements.

### ○ Minimum number of link levels

- The "more important" document the "closer" should be to the main page.
- But too many links on one page is not good either.

*Continued on next lecture...*