



ΑΕΙ ΠΕΙΡΑΙΑ Τ.Τ
ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΚΩΝ ΕΦΑΡΜΟΓΩΝ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΗΛΕΚΤΡΟΝΙΚΩΝ
ΥΠΟΛΟΓΙΣΤΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ Τ.Ε.

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

**Τεχνικές, μέθοδοι και κανόνες για τη βελτιστοποίηση ανάπτυξης
των ιστοσελίδων**

Δημητριάδης Ζαχαρίας

Εισηγητής: Ματιάτος Σπυρίδων, Καθηγητής Εφαρμογών

ΑΘΗΝΑ

ΑΠΡΙΛΙΟΣ 2016

Τεχνικές, μέθοδοι και κανόνες για τη βελτιστοποίηση ανάπτυξης ιστοσελίδων

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Τεχνικές, μέθοδοι και κανόνες για τη βελτιστοποίηση ανάπτυξης των ιστοσελίδων

Δημητριάδης Ζαχαρίας

A.M : 42080

Εισηγητής:

Ματιάτος Σπυρίδων, Καθηγητής Εφαρμογών

Εξεταστική Επιτροπή:

Ημερομηνία Εξέτασης //

ΕΥΧΑΡΙΣΤΙΕΣ

Η παρούσα πτυχιακή εργασία ολοκληρώθηκε μετά από επίμονες προσπάθειες. Θα ήθελα να ευχαριστήσω θερμά τον κ. Ματιάτο Σπυρίδωνα κυρίως για την εμπιστοσύνη που μου έδειξε και την υπομονή για τη βελτίωση της πτυχιακής με στόχο την υψηλότερη βαθμολογία.

Επίσης θέλω να απευθύνω τις ευχαριστίες μου στους γονείς μου, οι οποίοι στήριξαν τις σπουδές μου με διάφορους τρόπους, φροντίζοντας για την καλύτερη δυνατή μόρφωση μου.

ΠΕΡΙΛΗΨΗ

Σκοπός της παρούσας πτυχιακής είναι η μελέτη των τεχνικών, των μεθόδων και των κανόνων που πρέπει να ακολουθούνται για την βελτιστοποίηση της ανάπτυξης τόσο της δομής όσο και του περιεχομένου των ιστοσελίδων ώστε να επιτυγχάνεται όσο το δυνατόν υψηλότερη κατάταξη στα αποτελέσματα των Μηχανών Αναζήτησης (ΜΑ). Εφαρμογή των μεθόδων αυτών σε μια ιστοσελίδα αναμένεται να αυξήσει την επισκεψιμότητά της μέσω οργανικών αναζητήσεων (αποτελεσμάτων των ΜΑ στις αναζητήσεις με λέξεις κλειδιά - keywords), ενώ παράλληλα θα την κάνουν φιλική στον χρήστη. Θα καταγραφούν και θα αναλυθούν επίσης τα εργαλεία τα οποία χρησιμοποιούνται για τον ποιοτικό έλεγχο ανάπτυξης μιας ιστοσελίδας καθώς και την συμβολή τους στην βελτίωση της κατάταξής της στα αποτελέσματα των ΜΑ.

ABSTRACT

This thesis covers the techniques, methods and rules that must be followed to optimize the development of both structure and content of web pages in order to achieve as much as possible higher ranking in search engine results. Implementation of such methods in a web page is expected to increase its traffic via organic search, while it remains user friendly. We will record and analyze the tools which are used for checking the quality of a website as well as their contribution in page ranking.

ΕΠΙΣΤΗΜΟΝΙΚΗ ΠΕΡΙΟΧΗ: Βελτιστοποίηση ιστοσελίδων με βάση τους κανόνες SEO για τις διάφορες μηχανές αναζήτησης

Λέξεις Κλειδιά: βελτιστοποίηση ιστοσελίδων, μηχανή αναζήτησης, μέθοδοι- κανόνες seo, εργαλεία seo, seo για κινητές συσκευές

Key Words: page optimization, search engine, methods- rules of seo, seo for mobile users

ΠΕΡΙΕΧΟΜΕΝΑ

1 Μηχανές αναζήτησης και τρόποι λειτουργίας αυτών	9
1.1 Ορισμός μηχανών αναζήτησης	9
1.2 Ιστορική αναδρομή μηχανών αναζήτησης	10
1.3 Λειτουργία μηχανών αναζήτησης	20
1.3.1 Crawler- based μηχανές αναζήτησης	20
1.3.1.1 Web crawling, robot crawler(Αράχνη)	21
1.3.1.2 Web Indexing(Ευρετηρίαση)	24
1.3.1.3 Query processing(Επεξεργασία ερωτημάτων)	25
1.3.2 Human- powered μηχανές αναζήτησης	25
1.3.3 Meta- μηχανές αναζήτησης	26
1.3.4 Hybrid, Υβριδικές μηχανές αναζήτησης	27
2 Δημοφιλέστερες μηχανές αναζήτησης και λόγοι δημοφιλίας τους	29
2.1 M.A της Google	29
2.1.1 Από που προήλθε η ονομασία Google	30
2.1.2 Πολιτική της MA Google	30
2.1.3 Βασική τεχνολογία της Google, ο αλγόριθμος PageRank	32
2.2 M.A της Bing	33
2.2.1 Ονομασία της Bing	33
2.2.2 Feature της Bing που την ξεχωρίζουν από άλλες MA	34
2.3 M.A της Baidu	36
2.3.1 Ονομασία της Baidu	36
2.3.2 Λόγοι για τους οποίους η MA Baidu έχει ξεχωρίσει	37
2.4 Ποσοστά μεριδίου χρήσης που κατέχουν οι MA ανά τον κόσμο	38
2.5 Λόγοι δημοφιλίας M.A που οδήγησαν στη σημερινή διαμόρφωση	39
3 Μέθοδοι, κανόνες και τεχνικές βελτιστοποίησης	41
3.1 Θεμελιώδης αλλαγές SEO με την έλευση των Panda και Penguin	41
3.2 “Χρυσοί κανόνες” SEO	43
3.3 Μέθοδοι βελτιστοποίησης	44
3.3.1 White hat SEO	44
3.3.2 Black hat SEO	46
3.4 Εφαρμογή των white-hat seo τεχνικών	48

3.4.1 On- site SEO	49
3.4.2 Off- site SEO	57
3.5 Λέξεις κλειδιά και που στοχεύουν (Keywords, target group)	62
3.6 Μειονεκτήματα SEO	63
4 SEO tools- software	65
4.1 Εργαλεία αναζήτησης λέξεων κλειδιών	65
4.2 Εργαλεία ελέγχου ιστοσελίδας	69
4.3 Εργαλεία βελτιστοποίησης	73
4.4 Άλλα εργαλεία	75
5 SEO για χρήστες κινητών τηλεφώνων	77
5.1 Κατασκευή mobile friendly ιστοσελίδων σύμφωνα με τη google	77
5.1.1 Responsive web design	78
5.1.2 Δυναμική εξυπηρέτηση(Dynamic Serving)	81
5.1.2.1 Λειτουργία του vary http header	81
5.1.3 Ξεχωριστά URL	82
6 SEM- Marketing στις MA	87
6.1 Ορισμός SEM	87
6.1.1 PPC	88
6.2 Διαφορές SEO και PPC και κριτήρια επιλογής	89
7 Βιβλιογραφία	91

1 Μηχανές αναζήτησης και τρόποι λειτουργίας αυτών

Οι ΜΑ αποτελούν τη σύγχρονη πύλη εισόδου προς το διαδίκτυο. Το διαδίκτυο είναι το παγκόσμιο σύστημα διασυνδεδεμένων υπολογιστών κάτω από το οποίο λειτουργούν συγκεκριμένα πρωτόκολλα για να εξυπηρετούνται καθημερινά εκατομμύρια χρήστες σε όλο το κόσμο. Το διαδίκτυο αποτελεί το πιο σύγχρονο τρόπο ανταλλαγής πληροφοριών, δεδομένων (πολυμέσων), μηνυμάτων και χρησιμοποιείται ακόμα και για το εμπόριο. Ακόμα και παραδοσιακές μορφές τέχνης (όπως για παράδειγμα ο κινηματογράφος και η μουσική) μέσω της ψηφιακής τεχνολογίας παίρνουν την ίδια μορφή (αρχείων δεδομένων) με αντικείμενα που εκ πρώτης όψεως είναι εντελώς διαφορετικά (όπως για παράδειγμα η ιατρική επιστήμη ή κάποιο πρόγραμμα λογισμικού). Παρατηρείται λοιπόν μία συγκέντρωση γνώσης, ή αν είναι δυνατό να λεχθεί, πολιτιστικής κληρονομιάς, που σχετίζεται άμεσα με το Ίντερνετ. Σήμερα, ο όρος αυτός κατέληξε να αναφέρεται στο παγκόσμιο αυτό δίκτυο το οποίο σε συνδυασμό με την ολοένα αναπτυσσόμενη τεχνολογία έχει δημιουργήσει μια τεράστια αγορά από πληροφορίες. Το μεγάλο ερώτημα που προκύπτει πλέον είναι το "ποιος θα διοικήσει, ποιος θα ελέγξει την γνώση αυτή αλλά και πως όλη αυτή η πληροφορία διαμοιράζεται και γίνεται γνωστή στους χρήστες του"[1].

1.1 Ορισμός μηχανών αναζήτησης

Η ΜΑ είναι υπεύθυνη για το διαμοιρασμό και τον έλεγχο των πληροφοριών και των δεδομένων. Οι πληροφορίες αυτές είναι αποθηκευμένες σε διάφορες ιστοσελίδες του παγκόσμιου ιστού WWW. Μια ΜΑ είναι μια εφαρμογή που επιτρέπει την αναζήτηση κειμένων και αρχείων στο Διαδίκτυο. Αποτελείται από ένα πρόγραμμα υπολογιστή που βρίσκεται σε έναν ή περισσότερους υπολογιστές ή ακόμα και κινητές συσκευές στα οποία δημιουργεί μια βάση δεδομένων με τις πληροφορίες που συλλέγει από το διαδίκτυο. Το διαδραστικό περιβάλλον που εμφανίζεται στον τελικό χρήστη είναι το αποτέλεσμα χρησιμοποίησης της εφαρμογής και σύνδεσης του με κάποιο άλλον υπολογιστή συνδεδεμένο στο διαδίκτυο[2]. Για την επίτευξη της αναζήτησης πληροφοριών ο χρήστης απαιτεί να πληκτρολογήσει στο κατάλληλο πεδίο της εφαρμογής τις λέξεις που είναι σχετικές

είτε με τον τίτλο, είτε με την ιστοσελίδα είτε μια μικρή περιγραφή σχετική με το περιεχόμενο το οποίο θέλει να εμφανιστεί. Τέλος η εφαρμογή θα δείξει τα αποτελέσματα που έχει βρει με βάση κάποια κριτήρια. Σε περίπτωση που ο αριθμός των ιστοσελίδων που σχετίζονται με την αναζήτηση του χρήστη είναι μεγάλος, γίνεται η κατανομή αυτών σε μία ή περισσότερες σελίδες για την καλύτερη οργάνωση των αποτελεσμάτων. Αξίζει να αναφερθεί πως υπάρχουν πολλές MA οι οποίες όμως για την ίδια αναζήτηση μπορεί να επιστρέψουν διαφορετικά αποτελέσματα, συνεπώς η επιλογή της εφαρμογής εξαρτάται από τη κρίση του χρήστη.

1.2 Ιστορική αναδρομή μηχανών αναζήτησης

Archie

Η πρώτη MA ξεκίνησε από μία ομάδα φοιτητών και εθελοντών του πανεπιστημίου McGill (university school of computer science) του Μόντρεαλ το 1987. Πιο συγκεκριμένα ζητήθηκε από τους Peter Deatch, Emtage και Heelan να διασυνδέσουν το πανεπιστήμιο τους με το διαδίκτυο. Στα πρώιμα στάδια της υπηρεσίας Archie κατάφεραν να κρατήσουν σε λίστες (index) όλους τους δημόσιους ftp servers και να τις ενημερώνουν σε τακτά χρονικά διαστήματα, περίπου μία φορά το μήνα ώστε να μη σπαταλώνται και πολλοί πόροι από απομακρυσμένους διακομιστές. Οι καταχωρήσεις αυτές αποθηκεύονταν σε τοπικά αρχεία ώστε να αναζητηθούν χρησιμοποιώντας την εντολή grep του Unix. Τέλος η συγκεκριμένη υπηρεσία υπάρχει ακόμα και σήμερα, λειτουργεί όμως μόνο ως ιστοσελίδα, ως ένα δικτυακό μουσειακό έκθεμα :



Welcome to archie.icm.edu.pl

Archie Query Form

Search for: osarena

Database: Worldwide Anonymous FTP Polish Web Index

Search Type: Sub String Exact Regular Expression

Case: Insensitive Sensitive

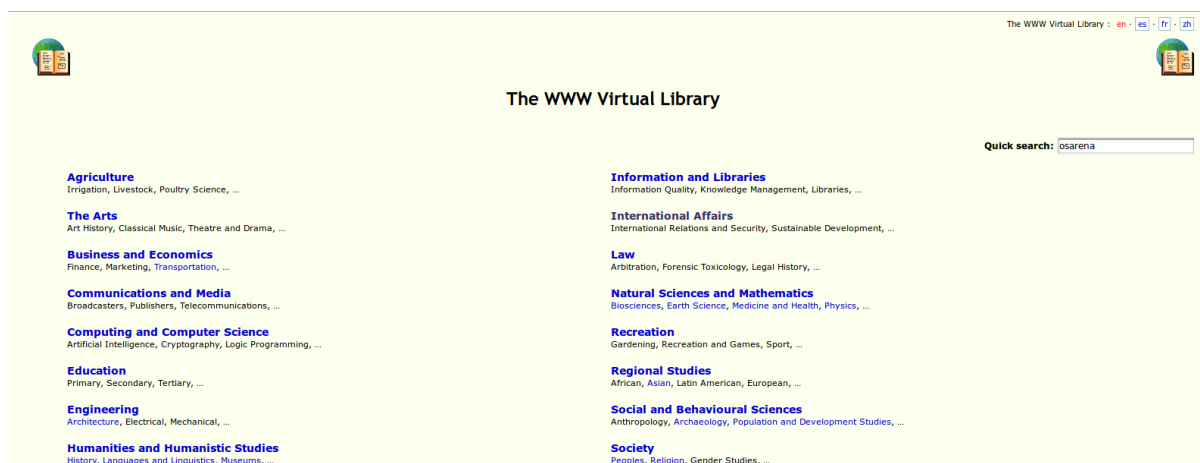
Do you want to look up strings only (no sites returned):
 NO YES

Output Format For Web Index Search: Keywords Only Excerpts Only Links Only

Εικόνα 1.1: Μηχανή αναζήτησης archie[3].

V-Lib

Λίγα χρόνια πιο μετά το 1992, εμφανίστηκε η MA V-Lib. Μια εικονική βιβλιοθήκη αναζήτησης από τον πατέρα του Διαδικτύου όπως θα χαρακτήριζαν πολλοί, τον Tim Berners-Lee η οποία φιλοξενείται στο Cern. Υπάρχει ακόμα και σήμερα, αλλά δεν μπορεί να χρησιμοποιηθεί ως μία γενική MA.



The WWW Virtual Library

Quick search: osarena

Agriculture
Irrigation, Livestock, Poultry Science, ...

The Arts
Art History, Classical Music, Theatre and Drama, ...

Business and Economics
Finance, Marketing, Transportation, ...

Communications and Media
Broadcasters, Publishers, Telecommunications, ...

Computing and Computer Science
Artificial Intelligence, Cryptography, Logic Programming, ...

Education
Primary, Secondary, Tertiary, ...

Engineering
Architecture, Electrical, Mechanical, ...

Humanities and Humanistic Studies
History, Languages and Linguistics, Museums, ...

Information and Libraries
Information Quality, Knowledge Management, Libraries, ...

International Affairs
International Relations and Security, Sustainable Development, ...

Law
Arbitration, Forensic Toxicology, Legal History, ...

Natural Sciences and Mathematics
Biosciences, Earth Science, Medicine and Health, Physics, ...

Recreation
Gardening, Recreation and Games, Sport, ...

Regional Studies
African, Asian, Latin American, European, ...

Social and Behavioural Sciences
Anthropology, Archaeology, Population and Development Studies, ...

Society
Peoples, Religion, Gender Studies, ...

Εικόνα 1.2: Μηχανή αναζήτησης V-Lib[3].

Ali-Web

Κάπου μέσα στο 1993, έκανε την εμφάνιση της η Ali-Web, στην οποία οι χρήστες μπορούσαν να καταχωρούν ιστοσελίδες. Υπάρχει και σήμερα σαν σελίδα. Μπορώντας να αποδώσει κάποια από τα καταγεγραμμένα αποτελέσματα και αυτή, κατατάσσεται στα μουσειακά είδη, κρατώντας μάλιστα μια εμφάνιση που παραπέμπει στην εποχή εκείνη:

Aliweb Home | September 11th, 2001

Aliweb osrena Search Reset

ADVANCED OPTIONS: Case Sensitive: Description URL Other

Display Fields Other Than Title? Description Keyword Other

Select Search Fields Title Description Keyword URL

Limit Results To: 20 matches Restrict to Domain: none

Search Type: Whole Words Record Type: Any

AllLinks - Links to the web's best sites all on one easy to navigate page!
The Web9000 Network makes it easy and affordable. [click here.](#)

Computing
Jump to Section: TOP Computing Entertainment Libria Money Newstand Recreation Research Shopping BOTTOM

Browsers Etc. Amaya | ArachneDoo | Chrome | Microsoft IE | Firefox & Thunderbird | Flashplayer | Maxthon | Netscape (old) | Opera | PDF Reader | Pegasus Mail | QuickTime | RealPlayer | Safari | Shockwave

CGI/Java CGI Resource | Extropia | Gamelan | Java Boutique | JavaScript | Matt's | Webscripts |

Companies AMD | Apple | Compaq | Cyrix | Dell | Digital | Electro Service | Gateway | HP | Intel | Winchip |

Free Services Announce: Addurl | ATC | Eneppomote | Shotgun Banners: Link Exchange | Smartclicks Chat: ActiveWorlds | Chatlist | Globe | Talk City | WBS | Yack! Email: Hotmail | Juno | MailExcite Hosting: Angelfire | Fortune City | Geocities | Trailerpark | Tripod | Web9000 | Xoom

Games A_Vault | Gamecenter | Game Finder | GamePao | GamePower | G_Domain | Games.net | GameSpot | Happy Puppy | OGR

Site Help Absolute | Builder.com | DJ.Dual | developer.com | Dev_Zone | HOME | HTMLgoodies | HTML_Guide | WebDeveloper | reallybig.com | Webmonkey | Webreference | Web_Review

Site Tools Color Chart | Cool Tool | Dr. HTML | FreeForm | GIF Wizard | NetMechanic | SBN Gallery | Site Garage

Software Dave Central | Download | Extreme Mac | Filemine | FilePile | Filaz | Freewaresplus | Freewareweb | Jumbo | Nonags | Softseek | Stroud's | Tucows | Windrivers | Winfiles | Winsite | ZDnet_Hoffiles

Support Newbie-U | Microsoft | PC-Help | Support_Help | Whalls

Entertainment
Jump to Section: TOP Computing Entertainment Libria Money Newstand Recreation Research Shopping BOTTOM

Movies Boxoffice | Cinemachine | Film.com | Film Scouts | In Theaters | IMDb | Moviefinder | MovieLink | Mr. Cranky | Real

Movie Studios Buena Vista | Disney | Dimension | Fox | Fox Searchlight | Hollywood | MGM/UA | Miramax | New Line | October | Paramount | Sony | Universal | Warner Bros

Music Ezines Allstar | Classical | CMJ Online | Ear.1 | MTV | RollingStone | SonicNet | VibeOnline

Music Concerts LiveConcerts | Live Online | Pollstar | Ticketmaster | WILMA

Music News Billboard | HitsWorld | Imusic | Newswire

Music Search IUMA | JazzCentral | LyricServer | MusicCentral | MusicSearch | URL | UnfURLed | WallOfSound

Stars Cyberlease | eDrive | EI Online | Ent. Asylum | E_Weekly | Engulter | Hollywood | Mr.Showbiz | People | Variety

TV click.TV | GIST | Prevue | Rock on TV | Total TV | TV Guide | UltimateTV

TV Networks ABC | CBS | Disney | FOX | ESPN | HBO | NBC | PBS | WB | UPN | SciFi

Living
Jump to Section: TOP Computing Entertainment Libria Money Newstand Recreation Research Shopping BOTTOM

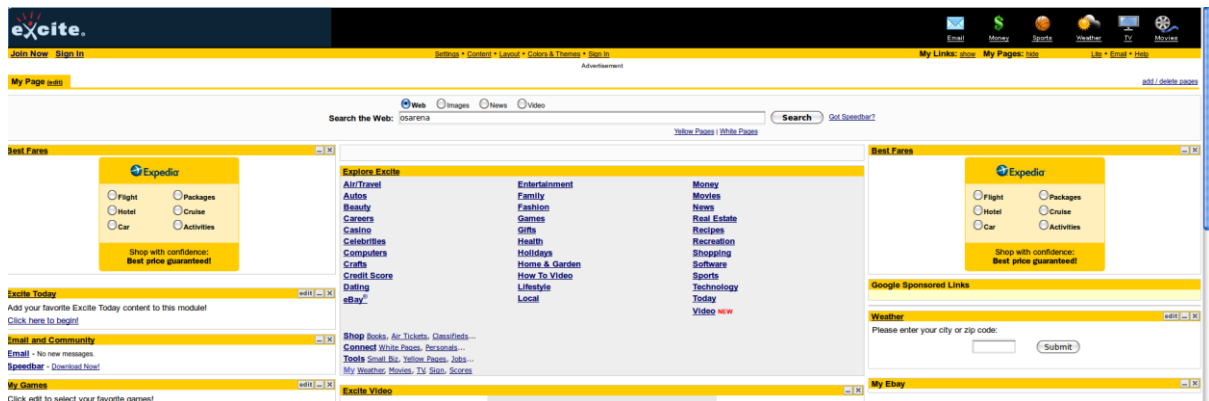
Art A.Resource | CultureFinder | Fine Site | The Guide | Playbill Buy/Sell | Doubtstake

Education .edu | A.S.D. | AskERIC | CollegeNet | Petersons | University | Web 66

Εικόνα 1.3: Μηχανή αναζήτησης Aliweb[3].

Excite

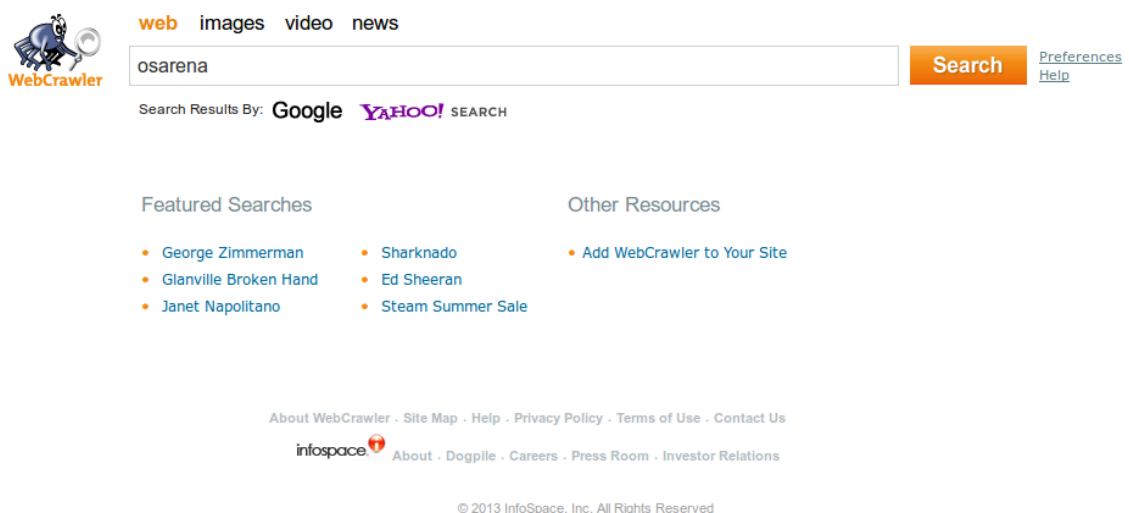
Το ίδιο ακριβώς διάστημα (1993), μετά από κάποιους μήνες αργότερα, εμφανίστηκε μια ακόμα MA, με πιο μοντέρνο interface. Το όνομα αυτής Excite, η οποία λόγω χρεωκοπίας πουλήθηκε το 2001. Σήμερα λειτουργεί κανονικά και πολλοί υποστηρίζουν πως πρόκειται για μια πολύ σπουδαία μηχανή. Αν και τα αποτελέσματα τα δίνει σε ένα καλό και καθαρό γραφικό περιβάλλον(interface) η κεντρική της σελίδα, θυμίζει αρκετά τα 90s:



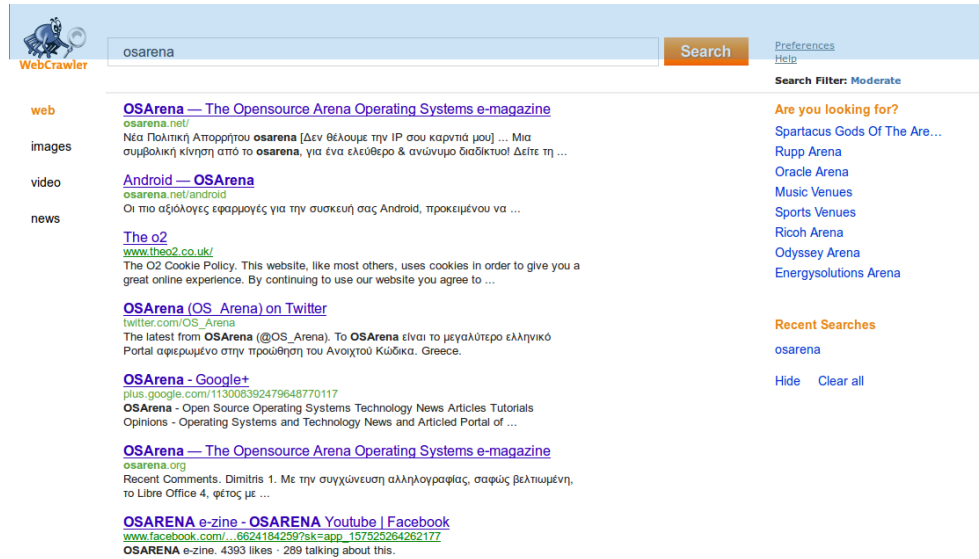
Εικόνα 1.4: Μηχανή αναζήτησης excite[3].

WebCrawler

Ένα χρόνο αργότερα, το 1994, τα πράγματα έχουν αρχίσει ν' αλλάζουν και κάνει την εμφάνιση της η WebCrawler, η πρώτη MA που τοποθετούσε σε ευρετήριο ολόκληρες τις σελίδες. Έγινε τόσο δημοφιλής που ορισμένες φορές, με το διαδίκτυο της εποχής ήταν σχεδόν αδύνατο να χρησιμοποιηθεί. Συνεχίζει να υπάρχει με μοντέρνα εμφάνιση, παρουσιάζοντας αποτελέσματα από τις μηχανές των Google και Yahoo:



Εικόνα 1.5: Μηχανή αναζήτησης Webcrawler[3].



Εικόνα 1.6: Αποτελέσματα μηχανής αναζήτησης webcrawler[3].

Yahoo

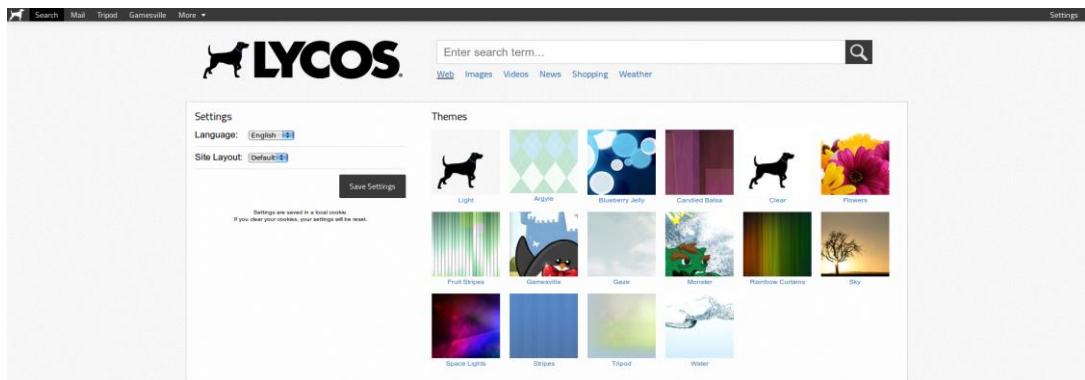
Την ίδια χρόνια ιδρύθηκε η yahoo από τους τότε τελειόφοιτους του πανεπιστημίου Στάνφορντ, *David Filo* και *Jerry Yang* οι οποίοι ήθελαν να ομαδοποιήσουν και να καταγράψουν διαδικτυακές τοποθεσίες μείζονος ενδιαφέροντος και να τις ταξινομήσουν σε θεματικές ενότητες. Αρχικά ο θεματικός κατάλογος ήταν δημοσιευμένος στον δικτυακό τόπο του Στάνφορντ (akebono.stanford.edu/yahoo) και στις 2 Μαρτίου 1995 εισήχθη στο χρηματιστήριο[4].



Εικόνα 1.7: Μηχανή αναζήτησης yahoo.

Lycos

Δεν περάσαν λίγοι μήνες και η MA που ξεχώρισε εκείνη την εποχή ήταν η Lycos, όπου δημοσιοποίησε έναν κατάλογο με πάνω από 54.000 εγγραφές, κατατάσσοντας τα αποτελέσματα σύμφωνα με την σχετικότητα ανάκτησης, ενώ συγχρόνως έκανε αντιστοίχιση προθέματος και εγγύτητας λέξης. Για την εμπορική εκμετάλλευση της μηχανής αυτής δημιουργήθηκε στα τέλη Ιουνίου του 1995 η εταιρεία Lycos Inc., ενώ τεχνολογικά αποτελεί έργο του Dr. Michael Mauldin στο Πανεπιστήμιο Carnegie Mellon. Ενώ ακόμα συνεχίζει να υπάρχει, περιλαμβάνει άνω του 90% των διευθύνσεων του Web!



Εικόνα 1.8: Μηχανή αναζήτησης lycos[3].

Altavista

Ένας θρύλος γεννιέται το 1995: η Altavista, η πρώτη μηχανή που επέτρεπε αναζητήσεις στην φυσική γλώσσα του καθενός, ενώ είχε απεριόριστο εύρος ζώνης, με προηγμένες τεχνικές αναζήτησης. Η προσθήκη και η διαγραφή κάποιου Url γινόταν μέσα σε 24 ώρες! Σήμερα δεν υπάρχει πλέον, αν και το url της (altavista.com) είναι ενεργό, κάνει redirect στην Yahoo η οποία και την απέκτησε αργότερα. Εν' τω μεταξύ, κάποιο διάστημα αργότερα έκανε την εμφάνιση της η Raging Search που απευθύνονταν σε αυτούς που δεν ήθελαν να βλέπουν το portal της Altavista παρέχοντας, έτσι ταχύτητα στην αναζήτηση. Και αυτή εξαγοράστηκε από την Yahoo, στην οποία κάνει redirect το url της. Έχοντας φτάσει στο τέλος του 1995, ξεπήδησε η Yahoo η οποία ήταν ο μεγάλος star, ιδιαίτερα την διετία 1995-1997. Ο Jerry Yang και ο David Filo, ήταν οι δημιουργοί

Τεχνικές, μέθοδοι και κανόνες για τη βελτιστοποίηση ανάπτυξης ιστοσελίδων

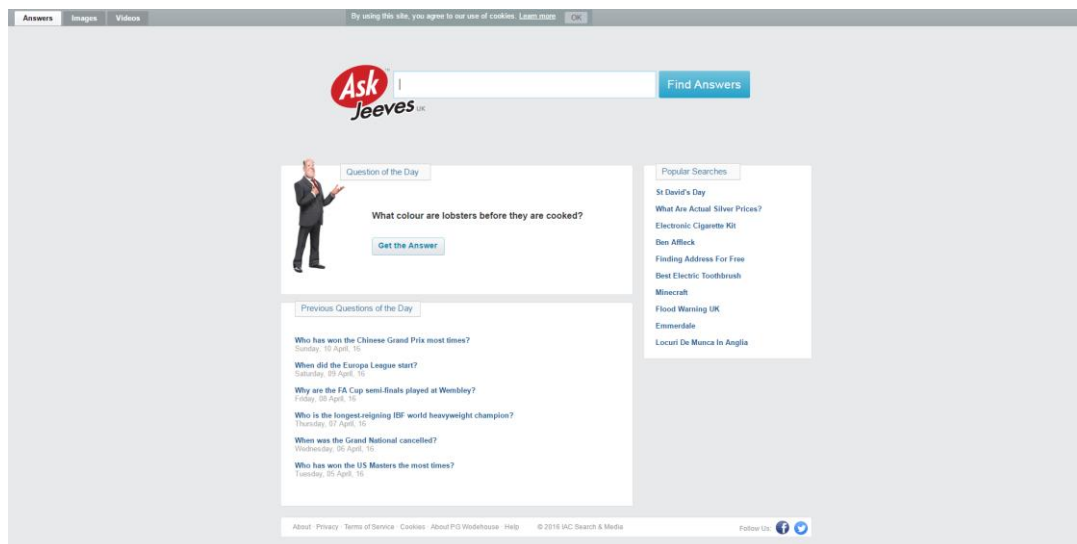
της και στην αρχή ήταν απλά μια συλλογή από αγαπημένες ιστοσελίδες. Με το μέγεθος της να αυξάνεται η ιδέα ήταν να τη μετατρέψουν σε MA. Η προσθήκη των url's, γινόταν χειροκίνητα και οι χρόνοι αναμονής εμφάνισης των αποτελεσμάτων ήταν μεγάλοι. Μια γεύση από τις τελευταίες εμφανίσεις:



Εικόνα 1.9: Μηχανή αναζήτησης altavista[3].

Ask Jeeves

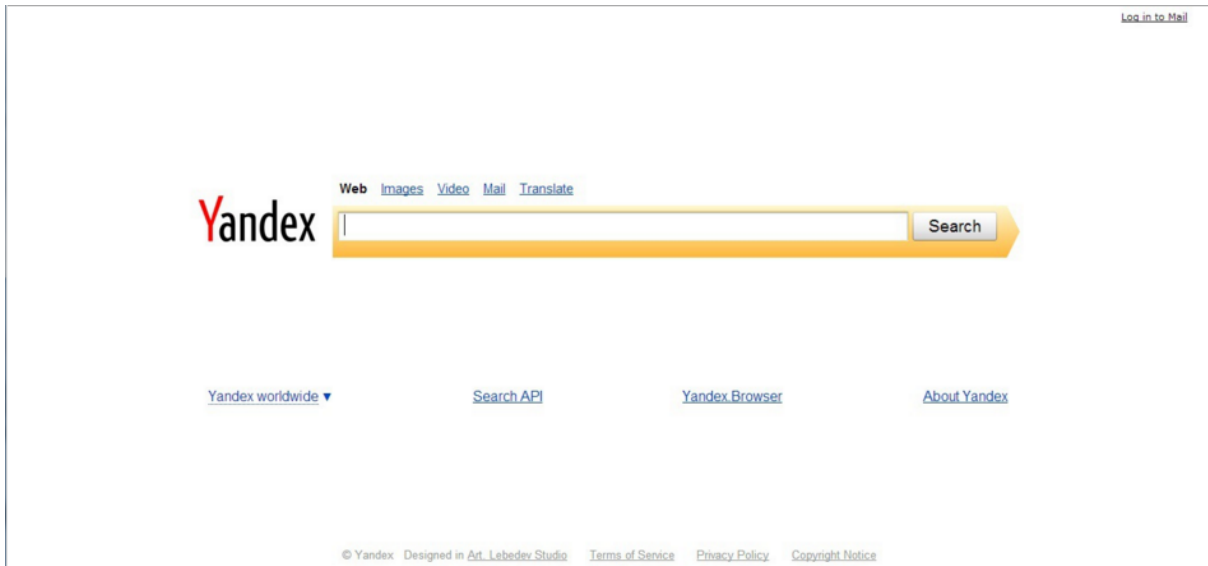
Κάπου εκεί, στα μέσα του 1997, έκανε την εμφάνιση της η Ask Jeeves. Η διαφορά της ήταν πως μπορούσε κάποιος να υποβάλει ερωτήματα στην φυσική του γλώσσα, με χρήστες να προσπαθούν να ταιριάξουν τα αποτελέσματα των αναζητήσεων. Άθελα της με αυτό τον τρόπο, συνέβαλε και στο spam. Υπάρχει ακόμα και σήμερα, με ένα μοντέρνο περιβάλλον, με αξιόλογα αποτελέσματα και δυνατότητα παραμετροποίησης και ελέγχου των cookies. Σήμερα έχει μετατραπεί σε Ask.com. Σήμερα το περιβάλλον αυτής είναι όπως φαίνεται στην εικόνα 1.10:



Εικόνα 1.10: Μηχανή αναζήτησης ask jeeves.

Yandex

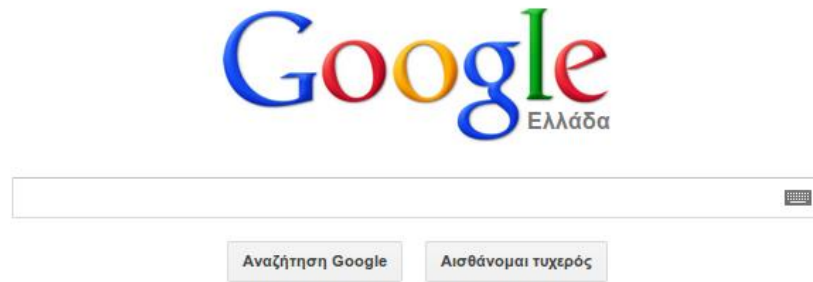
Την ίδια χρονιά ιδρύθηκε και πιο συγκεκριμένα το Σεπτέμβριο του 1997 η MA yandex και παρουσιάστηκε δημόσια στην softlool έκθεση της Μόσχας. Αρχικά είχε αναπτυχθεί από την comrtek και δραστηριοποιείται κυρίως στη Ρωσία[6].



Εικόνα 1.11: Μηχανή αναζήτησης yandex

Google

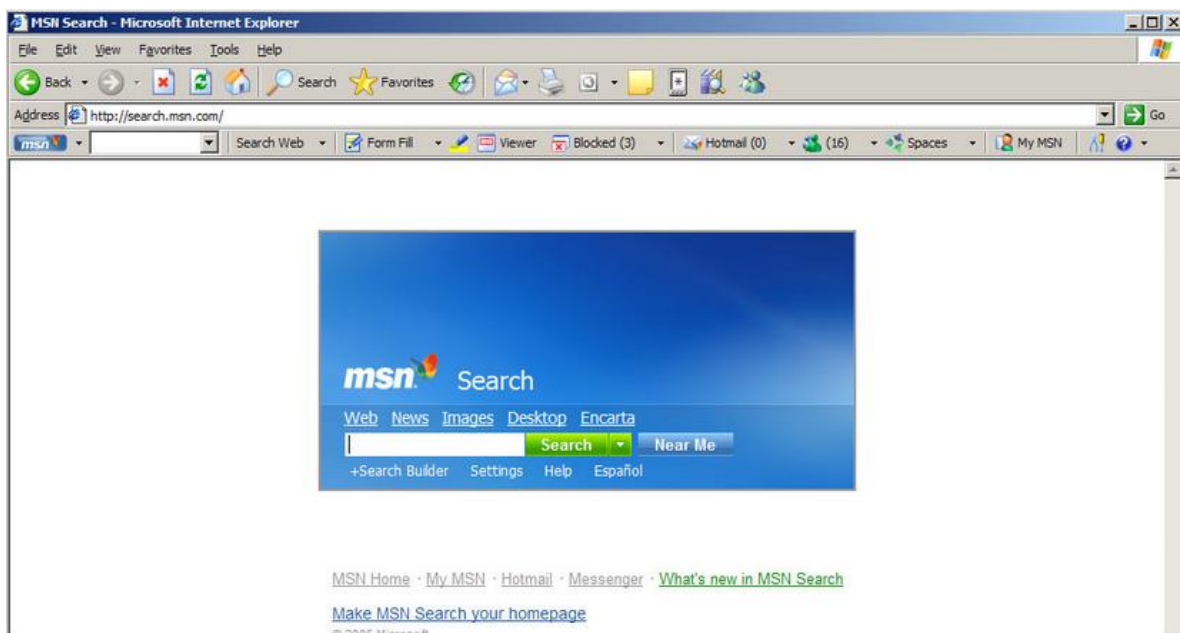
Το 1998 όμως έχει έρθει και ένα όνομα που αρκετοί το "λάτρεψαν" για πολλά χρόνια, η γνωστή σε όλους μας Google.. Οι νεαροί Larry Page και Sergey Brin, δημιούργησαν μια MA που αξιοποιούσε τις συνδέσεις και αναφορές από και προς άλλες ιστοσελίδες. Η Google συνεχίζει τόσες δεκαετίες μετά, έχοντας καταστεί ο γίγαντας και το σημείο αναφοράς στον χώρο (google it), με τον 97% (όπως λέγεται) του διαδικτύου, μα δεν είναι πλέον η ίδια αγαπημένη εταιρεία, όπως τα πρώτα χρόνια του ξεκινήματός της. Ένα χαρακτηριστικό και αμετάβλητο στα χρόνια, αποτελεί η κεντρική σελίδα της μηχανής της Google, όπου παρέμεινε λιτή και χρηστική:



Εικόνα 1.12: Μηχανή αναζήτησης google[3].

Msn-Bing

Το 1999 έκανε την εμφάνιση της και η MA Msn από τη μεγαλύτερη στα λειτουργικά συστήματα εκείνης της εποχής, microsoft. Αρκετά αργότερα, το 2004, αναβαθμίστηκε με νέους αλγορίθμους, ενώ μόλις έναν χρόνο αργότερα, απέρριπτε τα αποτελέσματα που βασίζονταν σε αλγορίθμους της Yahoo(από την οποία τελικά απέκτησε και ένα μέρος της, κάπου 6 με 7 χρόνια αργότερα). Το 2009 αλλάζει ονομασία και γίνεται bing.



Εικόνα 1.13: Μηχανή αναζήτησης msn[3].



Εικόνα 1.14: Μηχανή αναζήτησης bing.

Αξιοσημείωτες αναφορές

Αξίζει να αναφερθούμε και σε άλλες άγνωστες MA οι οποίες σήμερα μπορεί να μη μας θυμίζουν κάτι, όμως πριν από 15-20 χρόνια είχαν συμβάλει και αυτές με το τρόπο τους σε ότι έχουμε σήμερα, καθώς βοήθησαν στην ανάπτυξη τόσο του επιπέδου αναζήτησης αλλά και ευρετηρίασης ιστοσελίδων. Μερικές από αυτές είναι οι:

- Inktomi(το url της inktomi.berkeley.edu είναι ανενεργό),
- alltheweb/fast,
- hotbot (με ενεργό link hotbot.com),
- direct hit,
- looksmart(με ενεργό link search.looksmart.com),
- northern light,
- wisenut (με ενεργό link wisenut.com),
- goto/overture (όπου το link της κάνει redirect στη yahoo καθώς αγοράστηκε από αυτή),
- open directory-dmoz,
- infoseek,
- netscape/aol,
- northernlight,

- internetdatabase,
- deoji,
- netsearch,
- 1blink,
- aliweb,
- allonesearch,
- searchking,
- theinternetdirectories,
- 7search και
- highway61[3].

1.3 Λειτουργία μηχανών αναζήτησης

Όλες οι MA παρόλη τη διαφορετικότητα τους έχουν κάποιες κοινές λειτουργίες. Οι MA έχουν 2 βασικές λειτουργίες :crawling και indexing ώστε να παρέχουν στους χρήστες του διαδικτύου μία ταξινομημένη λίστα βάση τη κατάταξη των ιστοσελίδων. Με βάση τις παραπάνω λειτουργίες έχουν χτιστεί οι παρακάτω τύποι μηχανών αναζήτησης:

- Crawler-based MA
- Human- powered MA
- Hybrid MA
- Meta MA

1.3.1 Crawler- based μηχανές αναζήτησης

Αν φανταστούμε το παγκόσμιο δίκτυο WWW θα μπορούσαμε να το παρομοιάσουμε με ένα σταθμό με θεωρητικά άπειρες μοναδικές στάσεις(ιστοσελίδες) που η καθεμία έχει να δείξει κάτι διαφορετικό καθώς περιέχεται από διαφορετικό κείμενο,εικόνες και περιεχόμενο. Για το λόγο αυτό,οι MA αποτελούνται από 3 είδη λογισμικού ώστε να μας βοηθήσουν να περιηγηθούμε ευκολότερα. Τα οποία είναι :

- το spider software

- το index software
- το query software

1.3.1.1 Web crawling, robot crawler (Αράχνη)

Οι MA δεν ερευνούν σε πραγματικό χρόνο τον παγκόσμιο ιστό αλλά μία βάση δεδομένων που περιέχει κάποια αντίγραφα ιστοσελίδων. Οι ιστοσελίδες αυτές επιλέγονται ανάμεσα σε δισεκατομμύρια σελίδες στο ίντερνετ. Για την έρευνα αυτή χρησιμοποιούνται κάποια προγράμματα που ονομάζονται «ρομπότ» ή «αράχνες» (spider, crawler ή webcrawler). Το λογισμικό το οποίο σχετίζεται με το web crawling χρησιμοποιείται από τις MA για να σαρώσει και να συλλέξει πληροφορίες για τις ιστοσελίδες. Αυτές οι αράχνες ή spiders δεν συλλέγουν images, καταλαβαίνουν μονάχα το κείμενο των ιστοσελίδων, τα links και το URL. Συλλέγουν όση περισσότερη πληροφορία μπορούν στο συντομότερο δυνατό χρονικό διάστημα. Εφόσον τα spiders δεν συλλέγουν images δεν μπορούν να καταλάβουν οτιδήποτε είναι δημιουργημένο σε flash ή οτιδήποτε είναι εικόνα. Συνεπώς είναι σημαντικό να βεβαιωθούμε ότι σε όλες μας τις φωτογραφίες, images, flash, logos, videos κάνουμε χρήση του alt tag για να τα περιγράψουμε. Στην αντίθετη περίπτωση τα spiders θα τα αγνοήσουν και δεν θα δημιουργήσουν καμία προστιθέμενη αξία στη διαδικασία της προώθησης της ιστοσελίδας σας μέσω του search engine optimization το οποίο θα δούμε αργότερα[2],[7]. Επιπλέον αν παρατηρηθούν αλλαγές σε κάποιες ιστοσελίδες οι αράχνες αυτές θα τις καταλάβουν και θα βρουν αυτές τις αλλαγές με αποτέλεσμα να επηρεάσουν το τρόπο με τον οποίο είναι καταχωρημένες στις λίστες. Συνεπώς όλα παίζουν ρόλο και θα πρέπει να είμαστε ιδιαίτερα προσεκτικοί όπως θα δούμε και παρακάτω.

Για να αποφύγει μια αράχνη τον έλεγχο της ίδιας σελίδας ξανά και ξανά καταφεύγει συνήθως σε κάποιο είδος κανονικοποίησης συνδέσμων(url normalization). Υπάρχουν διάφοροι τρόποι κανονικοποίησης των url. Σύμφωνα με το standard URI **Generic Syntax** RFC 3986[8] που καθιερώθηκε τον Ιανουάριο του 2005 οι σωστοί τρόποι κανονικοποίησης περιλαμβάνουν:

Κανονικοποίηση και διατήρηση της σημασιολογίας του URL

- μετατροπή των url σε πεζούς χαρακτήρες.

`HTTP://www.Example.com/` → `http://www.example.com/`

Εικόνα 1.15: URL normalization σε πεζούς χαρακτήρες[9].

- μετατροπή χαρακτήρων σε κεφαλαία μετά από ακολουθίες διαφυγής

`http://www.example.com/a%c2%b1b` → `http://www.example.com/a%C2%B1b`

Εικόνα 1.16: Capitalizing letters σε χαρακτήρες διαφυγής[9].

- αποκωδικοποίηση των percent-encoded octets

`http://www.example.com/%7Eusername/` → `http://www.example.com/~username/`

Εικόνα 1.17: Αποκωδικοποίηση percent-encoded octets [9].

- αφαίρεση του προεπιλεγμένου port

`http://www.example.com:80/bar.html` → `http://www.example.com/bar.html`

Εικόνα 1.18: Αφαίρεση port[9].

Κανονικοποίηση που συνήθως διατηρείται η σημασιολογία του URL

- προσθήκη του slash

`http://www.example.com/alice` → `http://www.example.com/alice/`

Εικόνα 1.19: Προσθήκη slash στο τέλος του url[9].

- και αφαίρεση των τελειών (“.”)

`http://www.example.com/./a/b/./c/./d.html` → `http://www.example.com/a/c/d.html`

Εικόνα 1.20: Αφαίρεση “.”[9].

Κανονικοποίηση όπου τη σημασιολογία του URL δεν διατηρείται

- αφαίρεση του index ,fragments και των διπλότυπων slashes

`http://www.example.com/default.asp` → `http://www.example.com/`
`http://www.example.com/bar.html#section1` → `http://www.example.com/bar.html`
`http://www.example.com/foo//bar.html` → `http://www.example.com/foo/bar.html`

Εικόνα 1.21: Αφαίρεση του index, του fragment και των διπλών slashes του url αντίστοιχα[9].

- αντικατάσταση της IP με το domain name

`http://208.77.188.166/` → `http://www.example.com/`

Εικόνα 1.22: Αντικατάσταση της IP[9].

- περιορισμός του ονόματος των πρωτοκόλλων στο url

`https://www.example.com/` → `http://www.example.com/`

Εικόνα 1.23: Περιορισμός του ονόματος πρωτοκόλλου στο url[9].

- αφαίρεση ή προσθήκη του “www”

`http://www.example.com/` → `http://example.com/`

Εικόνα 1.24: Αφαίρεση του “www” από το url[9].

- ταξινόμηση των query

`http://www.example.com/display?lang=en&article=fred` → `http://www.example.com/display?article=fred&lang=en`

Εικόνα 1.25: Ταξινόμηση των query[9].

- αφαίρεση μη χρησιμοποιημένων μεταβλητών των query και “?” όταν το query είναι κενό

`http://www.example.com/display?id=&sort=ascending` → `http://www.example.com/display`

`http://www.example.com/display?` → `http://www.example.com/display`

Εικόνα 1.26: Αφαίρεση μη χρησιμοποιημένων μεταβλητών και “?” στο query αντίστοιχα[9].

Όπως καταλαβαίνετε το κόστος ενός λογισμικού αραχνών(spider software) είναι πολύ υψηλό. Συνεπώς για να διατηρείται σε σχετικά κανονικά επίπεδα ακολουθούνται κάποιοι μέθοδοι όπως:

- Περιορισμός των ιστοσελίδων που υπάρχουν στη βάση δεδομένων, διαγράφοντας τις παλιότερες ιστοσελίδες.
- Περιορισμός του αριθμού των επισκέψεων της αράχνης σε ίδιες σελίδες.
- Περιορισμός επίσκεψης της αράχνης σε κάποιες περιοχές του διαδικτύου.

Επίσης πρέπει να αναφερθεί πως παρόλο που τα λογισμικά των αραχνών έχουν την ίδια λειτουργία, ο αλγόριθμος προγραμματισμού της κάθε αράχνης είναι διαφορετικός. Οι κατηγορίες των αραχνών διακρίνονται ανάλογα με τον αλγόριθμο υπολογισμού του βαθμού βαρύτητας(αλγόριθμος Page Rank βλ. 2.1.3). Μερικοί από αυτούς είναι sharksearch, info spiders και άλλοι[45].

1.3.1.2 Web Indexing (Ευρετηρίαση)

Αφού τελειώσει με τη σάρωση(crawling) των ιστοσελίδων η MA δημιουργεί ένα κατάλογο ή ευρετήριο. Η διαδικασία αυτή ονομάζεται ευρετηρίαση ή αλλιώς indexing. Ο αλγόριθμος που καλύπτει αυτή τη διαδικασία ευρετηριάζει τις λέξεις ή ζευγάρια λέξεων και φράσεων ώστε να καταλάβει τη θεματολογία στην οποία αναφέρεται η κάθε ιστοσελίδα. Στη συνέχεια αξιολογεί τη πληρότητα και τη πληροφόρηση που παρέχουν οι διάφορες ιστοσελίδες και τις αποθηκεύει ανάλογα με το πόσο σημαντικές τις θεωρεί.

Ο σκοπός της αποθήκευσης ενός index είναι για τη βελτιστοποίηση της επίδοσης και της ταχύτητας της MA ώστε να εξυπηρετεί γρηγορότερα το χρήστη όταν του στέλνει ένα ερώτημα. Χωρίς ευρετήριο, η MA θα αναγκαζόταν να

ξανακάνει σαρώσεις, κάτι το οποίο θα χρειαζόταν πολύ χρόνο και μεγάλη υπολογιστική ισχύ. Για παράδειγμα αν είχαμε ένα index με 10.000 έγγραφα μπορούμε να πούμε ότι για την αναζήτηση θα χρειαστούμε μερικά milliseconds, ενώ για μια διαδοχική αναζήτηση 10.000 index θα χρειαζόμασταν κάποιες ώρες[10].

1.3.1.3 Query processing (Επεξεργασία ερωτημάτων)

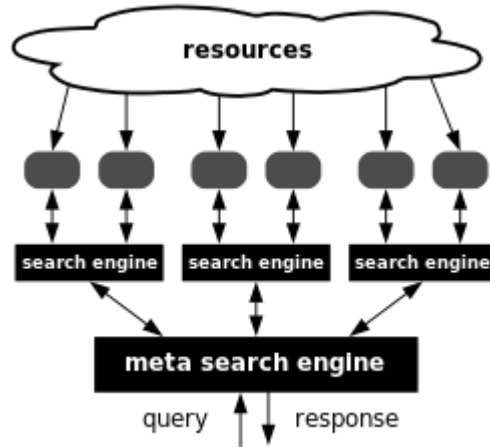
Είναι αυτό που βλέπουμε όταν χρησιμοποιούμε τις MA. Πληκτρολογώντας μία λέξη(keyword) στο κατάλληλο πεδίο της MA, μας εμφανίζονται τα οργανικά αποτελέσματα(βλ. κεφάλαιο 6.1) για αυτό. Για να εμφανιστούν αυτά τα αποτελέσματα, το query software έχει ψάξει στις ευρετηριασμένες σελίδες που έχουν αποθηκευτεί στις MA και παρουσιάζει αυτά τα οποία θεωρεί τα πιο σχετικά και σημαντικά για το keyword που έχει χρησιμοποιηθεί. Οι MA λοιπόν συλλέγουν, αναλύουν, ευρετηριάζουν, αξιολογούν και αποθηκεύουν τις ιστοσελίδες που υπάρχουν στο διαδίκτυο και διαθέτουν το αποτέλεσμα αυτών των εργασιών σε όσους πραγματοποιούν αναζητήσεις με συγκεκριμένα keywords. Στην ουσία οι MA ψάχνουν στο ευρετήριό τους όλα εκείνα τα στοιχεία που καλύπτουν το ερώτημα που έθεσε ο χρήστης και παρουσιάζει μια λίστα με τα αποτελέσματα-απαντήσεις του ερωτήματος. Αξίζει να σημειωθεί ότι διαφορετικές MA επιστρέφουν διαφορετικά αποτελέσματα καθώς οι αράχνες είναι διαφορετικά οργανωμένες για την κάθε MA [7].

1.3.2 Human- powered μηχανές αναζήτησης

Οι συγκεκριμένες MA βασίζονται κατά κύριο λόγο στον άνθρωπο. Αυτό συμβαίνει διότι η καταχώρηση των ιστοσελίδων πρέπει να γίνει είτε από τον κάτοχο της, είτε από τους συντάκτες του αντίστοιχου καταλόγου που κάνουν την αξιολόγηση αυτών ενώ συνήθως γίνεται επι πληρωμή. Ένα ερώτημα αναζητείται μόνο με βάση τη περιγραφή η οποία έχει καταχωρηθεί. Οποιοσδήποτε είδους βελτιστοποιήσεις μιας σελίδας η οποία βρίσκεται σε τέτοιου είδους κατάλογο δεν επιφέρει καμία απολύτως αλλαγή. Εν αντιθέτως είναι πιθανό μια πολύ καλή ιστοσελίδα με αξιόλογο περιεχόμενο να καταχωρηθεί δωρεάν από τους συντάκτες χωρίς ο κάτοχος η διαχειριστής της να παρουσιάσει κάποιο ενδιαφέρον και να προτείνει τη καταχώρηση της[12],[13].

1.3.3 Meta- μηχανές αναζήτησης

Σε αντίθεση με τις απλές MA, οι οποίες χρησιμοποιούν μια αράχνη για να συγκεντρώσουν μια δική τους βάση δεδομένων, οι μεταμηχανές δεν διαθέτουν δικό τους ευρετήριο, αλλά αντλούν τα αποτελέσματά τους από τα ευρετήρια άλλων MA. Οι μεταμηχανές στέλνουν τις λέξεις- κλειδιά ταυτόχρονα σε μία σειρά από προκαθορισμένες MA(ή και θεματικούς καταλόγους) και παρουσιάζουν ένα μέρος από τα αποτελέσματα της κάθε μίας. Καθώς ο μηχανισμός των μεταμηχανών παραμένει λίγο χρόνο σε κάθε βάση δεδομένων ανακτά συχνά μόνο το 10% των αποτελεσμάτων από κάθε βάση. Η πλειοψηφία των μεταμηχανών δεν παρέχουν πάνω 10 συνδέσμους από τη κάθε μηχανή και γενικά δεν επικοινωνούν με τις μεγαλύτερες. Συνήθως τα αποτελέσματα τα οποία έχουν προτεραιότητα και εμφανίζονται πρώτα είναι από διαφημίσεις[46]. Σε γενικές γραμμές η χρήση των μεταμηχανών ενδείκνυται για απλές αναζητήσεις (ένας όρος ή μία φράση), για να διαπιστώσουμε αν κάποιος όρος έχει τα αναμενόμενα αποτελέσματα ή για να εξοικονομήσουμε χρόνο αποφεύγοντας να επισκεφτούμε διαδοχικά πολλές μηχανές. Από την άλλη αν μία αναζήτηση περιέχει περισσότερες λέξεις- κλειδιά ή έχει πολύπλοκη λογική (λογική των τελεστών) είναι πολύ πιθανό να χαθεί ένα μεγάλο μέρος των αποτελεσμάτων, γιατί δεν έχουν όλες οι MA το ίδιο πρωτόκολλο αναζήτησης, δεν υποστηρίζουν δηλαδή όλες οι μηχανές την ίδια λογική εισαγωγής όρων. Βασικό πλεονέκτημα της συγκεκριμένης μηχανής θα μπορούσα να πει κανείς πως είναι η δυνατότητα που προσφέρεται στο χρήστη να αναζητεί μια μία λέξη σε διάφορες MA εξοικονομώντας χρόνο σε περίπτωση που αναλάμβανε ο ίδιος να υποβάλλει το ερώτημα για περισσότερες από μία MA. Η ίδια η MA αναλαμβάνει να υποβάλλει το ερώτημα του χρήστη σε διάφορες MA με τη κατάλληλη μορφή ώστε να πραγματοποιηθεί η αναζήτηση χωρίς προβλήματα[11].



Εικόνα 1.27: Υποβολή ερωτήματος από μία meta-μηχανή σε άλλες MA.

1.3.4 Hybrid, Υβριδικές μηχανές αναζήτησης

Μία υβριδική MA ή HSE (συντομογραφία του Hybrid Search Engine) χρησιμοποιεί ένα συνδυασμό από crawler-based αποτελέσματα και human-based αποτελέσματα. Με λίγα λόγια, μια τέτοια MA δίνει τη δυνατότητα στον ενδιαφερόμενο διαχειριστή ιστοσελίδας να καταχωρήσει την ιστοσελίδα του στον κατάλογο και παράλληλα αφήνει το αυτοματοποιημένο πρόγραμμα της μηχανής αναζήτησης να εισχωρήσει στο περιεχόμενο της ιστοσελίδας και να το αποθηκεύσει. Οι σημερινές MA τείνουν προς αυτή τη κατεύθυνση με τις πιο γνωστές από αυτές τις: Yahoo και Google[14].

2 Δημοφιλέστερες μηχανές αναζήτησης και λόγοι δημοφιλίας τους

Στο συγκεκριμένο κεφάλαιο θα δούμε τις δημοφιλέστερες MA, πως ξεκίνησαν και εξελίχθηκαν. Επίσης θα δούμε και τους λόγους που αυτές είναι οι πιο διαδεδομένες και τι ποσοστά κατέχουν στη παγκόσμια αγορά σύμφωνα με αναλύσεις.

2.1 M.A της Google

Στη πρώτη θέση της πιο δημοφιλούς MA δε θα μπορούσε να είναι άλλη από τη γνωστή σε όλους μας google. Οι ιδρυτές αυτής Larry Page και Sergey Brin, είχαν ως στόχο την ανάπτυξη μιας MA άλλα με διαφορετική προσέγγιση στο τρόπο με τον οποίο γινόταν η ηλεκτρονική αναζήτηση. Όπως είδαμε και προηγουμένως μέχρι τη δημιουργία της google υπήρχαν και άλλες MA. Βασικές όμως κατηγορίες MA που συναντούσαμε μέχρι εκείνη τη στιγμή ήταν πρώτα αυτές με τη μορφή ευρετηρίου οι οποίες συντηρούνταν από ανθρώπους και κατέτασσαν χειροκίνητα τις διάφορες ιστοσελίδες ανάλογα με τη θεματολογία τους. Αυτό είχε ως αποτέλεσμα μεγάλο κόστος δημιουργίας και συντήρησης ενώ η εξέλιξη τους γινόταν σε πολύ αργούς ρυθμούς. Στη δεύτερη κατηγορία συναντήσαμε τις αυτοματοποιημένες MA οι οποίες δημιουργούσαν και συντηρούσαν αυτόματα ευρετήρια ιστοσελίδων με βάση τις λέξεις κλειδιά ή τα μετα-δεδομένα που είχαν οι ιστοσελίδες. Εκτός από το ότι αυτή η μέθοδος έδινε χαμηλής ποιότητας αποτελέσματα στις αναζητήσεις των χρηστών, ήταν πολύ εύκολο να παραπλανηθούν οι συγκεκριμένες MA κάνοντας μια σελίδα να φαίνεται πρώτη στα αποτελέσματα.

Η ίδια ήταν η πρώτη που έκανε μεγάλη χρήση του υπερκειμένου (hypertext) των ιστοσελίδων για την κατάταξη τους. Με βάση αυτό το τρόπο αξιολογούσε τις σελίδες με τα εξής κριτήρια, τις λέξεις κλειδιά αλλά και το πόσο γνωστή είναι στο διαδίκτυο. Με τη φράση πόσο γνωστή ήταν στο διαδίκτυο εννοούμε όχι πόσο γνωστή ήταν στους χρήστες αλλά σε άλλες ιστοσελίδες. Με το τρόπο αυτό αξιολογούταν η εγκυρότητα των σελίδων και παρέχονταν στους χρήστες πιο αξιόπιστα αποτελέσματα. Τέλος, επειδή όλες αυτές οι διαδικασίες

είναι αυτοματοποιημένες το κόσμος συντήρησης του ευρετηρίου ήταν πολύ χαμηλό[16].

2.1.1 Από που προήλθε η ονομασία Google

Αρχικά η ονομασία που είχε δοθεί από τους Page και Brin ήταν Backrub, επειδή το σύστημα έλεγχε τα backlink για να εξετάσει τη σημαντικότητα των ιστοσελίδων. Τελικά όμως δόθηκε το όνομα google. Η ονομασία «Google» προήλθε από αναγραμματισμό της λέξης «Googol», τον μαθηματικό όρο που εισήγαγε ο Μίλτον Σιρότα για τον αριθμό 10^{100} που γράφεται: «1 ακολουθούμενο από 100 μηδενικά». Με τον όρο αυτόν η Google επιθυμεί να υποδηλώσει την αποστολή της εταιρίας να οργανώσει το τεράστιο πλήθος πληροφοριών του Ίντερνετ. Στην αρχή το Google έτρεχε στην ιστοσελίδα του Πανεπιστημίου Στάνφορντ, με τα domain google.stanford.edu και z.stanford.edu. Το όνομα domain για την Google κατοχυρώθηκε στις 15 Σεπτεμβρίου του 1997 και δημιουργήθηκε η εταιρία στις 4 Σεπτεμβρίου του 1998[16].

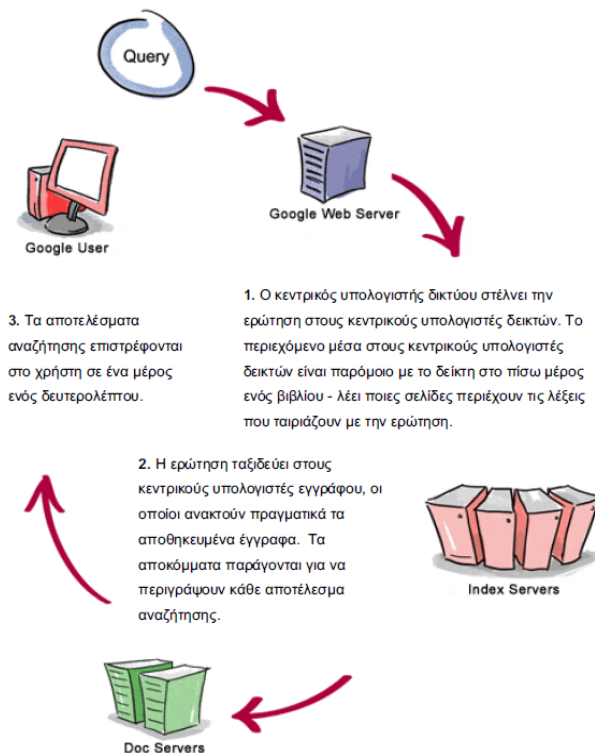


Εικόνα 2.1: Λογότυπο google[17].

2.1.2 Πολιτική της MA Google

Η Google εστίασε τις προσπάθειές της στην ανάπτυξη της "τέλειας μηχανής αναζήτησης", που όπως την ορίζει ο Larry Page, συνιδρυτής της Google, "πρέπει να καταλαβαίνει επακριβώς τι εννοεί ο χρήστης και πρέπει να δίνει επακριβώς αυτό που θέλει". Το σημαντικότερο θέμα που είχαν να επιλύσουν οι δημιουργοί

της, ήταν αυτό της ποιότητας των αποτελεσμάτων που φέρνει μια ΜΑ, στις αναζητήσεις των χρηστών του διαδικτύου. Για το σκοπό αυτό, η Google κυνήγησε με σθένος την καινοτομία και αρνήθηκε να δεχθεί τους περιορισμούς των υπάρχοντων μοντέλων. Αξίζει να σημειωθεί πως το 1997, μόνο μια από τις τέσσερις καλύτερες εμπορικές ΜΑ μπορούσε να βρει τον εαυτό της (να φέρει την αρχική ιστοσελίδα της, στα πρώτα δέκα αποτελέσματα, σε μία αναζήτηση του ονόματός της). □Ως αποτέλεσμα, η Google ανάπτυξε τη δική της υποδομή εξυπηρέτησης και την πρωτοποριακή τεχνολογία PageRank™, έναν αλγόριθμο σπουδαιότητας ιστοσελίδων που άλλαξε τον τρόπο διεξαγωγής των αναζητήσεων. Αυτή η καινοτομία είχε αποτελέσματα, καθώς εξασφάλισε πιο γρήγορους χρόνους απόκρισης, μεγαλύτερη δυνατότητα κλιμάκωσης και χαμηλότερο κόστος. Πρόκειται για μια ιδέα που βρήκε μιμητές, ενώ η Google συνεχίζει να βελτιώνει την τεχνολογία back-end για ακόμα καλύτερες επιδόσεις. Αξίζει επίσης να αναφερθεί πως η διάρκεια ζωής μιας ερώτησης στο google διαρκεί λιγότερο από μισό δευτερόλεπτο και περιλαμβάνει διάφορα στάδια πριν τη παράδοση της τελικής πληροφορίας στο χρήστη.



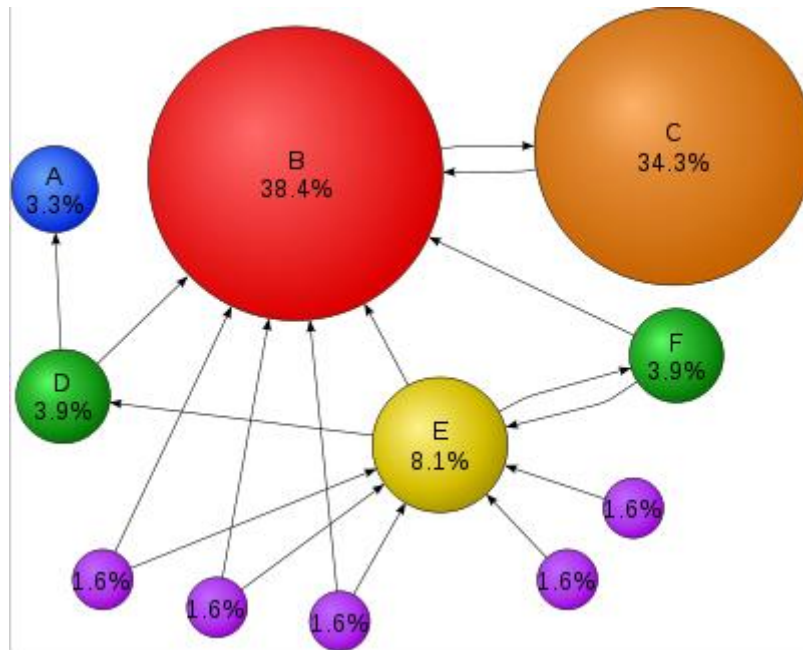
Εικόνα 2.2: Στάδια ενός query.

2.1.3 Βασική τεχνολογία της Google, ο αλγόριθμος PageRank

Ο πρώτος και βασικός αλγόριθμος ο οποίος χρησιμοποίησε η google για να ταξινομήσει τα αποτελέσματα της μηχανής της και για τον οποίο έχει γίνει γνωστή είναι ο PageRank. Ο συγκεκριμένος αλγόριθμος ονομάστηκε έτσι από το Larry Page. Χρησιμοποιείται για να καταγράψει τη σημαντικότητα μιας σελίδας σαρώνοντας τον αριθμό των συνδέσμων που έχει αυτή σε ξένες σελίδες. Η υπόθεση στην οποία στηρίχτηκε, ήταν πως οι σημαντικές ιστοσελίδες είναι πιθανό να αναφέρονται και σε άλλους διαδικτυακούς τόπους. Ο συγκεκριμένος αλγόριθμος αποτελεί έναν αλγόριθμο ανάλυσης των συνδέσμων των ιστοσελίδων όπου για κάθε στοιχείο αποδίδει ένα αριθμό βαρύτητας. Οι μετρήσεις αυτές αποσκοπούν στη μέτρηση της σημαντικότητας της σελίδας και το εκχωρημένο “βάρος” που έχει δοθεί σε κάθε στοιχείο αναφέρεται στο PageRank και συμβολίζεται με PR. Το PR μίας σελίδας ορίζεται αναδρομικά και εξαρτάται από τον αριθμό των σελίδων αλλά και το PR αυτών. Αν η διεύθυνση μιας σελίδας εντοπίζεται σε πολλές άλλες ιστοσελίδες με υψηλό PR, τότε αυτόματα αυτή έχει υψηλή κατάταξη. Για περισσότερες πληροφορίες σχετικά με το μαθηματικό κομμάτι του PR μπορείτε να βρείτε στη πηγή [19].

Παράδειγμα

Έστω ότι βρισκόμαστε στο πιο απλό δίκτυο που μπορούμε να αναπαραστήσουμε και το PR εκφράζεται με τη μορφή ποσοστών για χάρη του παραδείγματος. Παρατηρώντας τη φωτογραφία 2.3 θα παρατηρήσουμε πως η σελίδα C έχει μεγαλύτερο PR σε σχέση με τη σελίδα E, παρόλο που αυτή έχει περισσότερα εισερχόμενα links προς αυτή. Η σελίδα C μπορεί να περιέχει μόνο ένα εισερχόμενο link, όμως αυτό έρχεται από τη σελίδα B η οποία έχει υψηλό PR, πράγμα που σημαίνει πως αυτό που δείχνει είναι έγκυρο και αξιόπιστο αυξάνοντας αυτόματα το PR του C.



Εικόνα 2.3: Διαφορετικές ιστοσελίδες με το PR τους [19].

2.2 M.A της Bing

Η Bing αποτελεί μια MA που ανήκει στη microsoft. Οι υπηρεσίες που προσφέρει βασίζονται σε προηγούμενες MA όπως οι : msn search, windows live search και αργότερα της live search. Η Bing προσφέρει μία πληθώρα από αναζητήσεις, συμπεριλαμβανομένων των web, video, image και map. Για τη λειτουργία της έχει χρησιμοποιηθεί η ASP.NET γλώσσα προγραμματισμού και ακολουθεί τις αρχές σχεδίασης metro της microsoft.

2.2.1 Ονομασία της Bing

Για την επιλογή του ονόματος της MA η microsoft αποφάσισε πως θα έπρεπε να είναι κάποιο το οποίο θα το θυμόταν εύκολα ο χρήστης ενώ παράλληλα θα ήταν σύντομο και θα μπορούσε να λειτουργήσει ως μια διεύθυνση url παγκοσμίως. Τελικά βοηθήθηκε από μια ομάδα συμβούλων για την αναζήτηση του καλύτερου δυνατού ονόματος της νέας MA. Το όνομα Bing θυμίζει τη λέξη “bingo” καθώς χρησιμοποιείται για να σημάνει πως κάτι έχει βρεθεί ή πραγματοποιηθεί(παιχνίδι Bingo). Ενώ ακόμα δοκιμαζόταν εσωτερικά η λειτουργία της από τους υπαλλήλους της microsoft, αυτοί χρησιμοποιούσαν το όνομα kumo το οποίο προέρχεται από την ιαπωνική λέξη αράχνη θέλοντας να

Τεχνικές, μέθοδοι και κανόνες για τη βελτιστοποίηση ανάπτυξης ιστοσελίδων

αναφερθούν στο τρόπο με τον οποίο οι αράχνες- crawlers των MA συμπεριφέρονται στο διαδίκτυο. Ο πρόεδρος της microsoft online services Qi Lu ανακοίνωσε ότι το όνομα στα κινέζικα θα είναι bì yǐng το οποίο σημαίνει “πολύ βέβαιο ότι θα ανταποκριθούν” ή “πολύ βέβαιο ότι θα απαντήσει”, θέλοντας να πει φυσικά πως το ερώτημα του χρήστη θα απαντηθεί[22].



Εικόνα 2.4: Λογότυπο της Bing[22].

2.2.2 Feature της Bing που την ξεχωρίζουν από άλλες MA

Η MA της Bing προσφέρει λειτουργίες αναζήτησης που σίγουρα δεν έχουμε συναντήσει σε άλλες, για τις οποίες την κάνει σίγουρα να ξεχωρίσει. Αυτές είναι οι:

Flight cost prediction: Σίγουρα ένα από τα καλύτερα χαρακτηριστικά της, καθώς έχει την ικανότητα να προβλέπει και να δείχνει τις τιμές των εισιτηρίων για διάφορες πτήσεις τις οποίες ενδιαφέρεται για να πετάξει κάποιος. Με ειδικά χρωματικούς δείκτες (κόκκινο και πράσινο) υποδεικνύει στο χρήστη για την αυξομείωση της τιμής καθώς και τις ημερομηνίες. Συνεπώς μπορεί να σε βοηθήσει για την επιλογή της ημερομηνίας του ταξιδιού για να έχεις κάποιο καλύτερη τιμή.

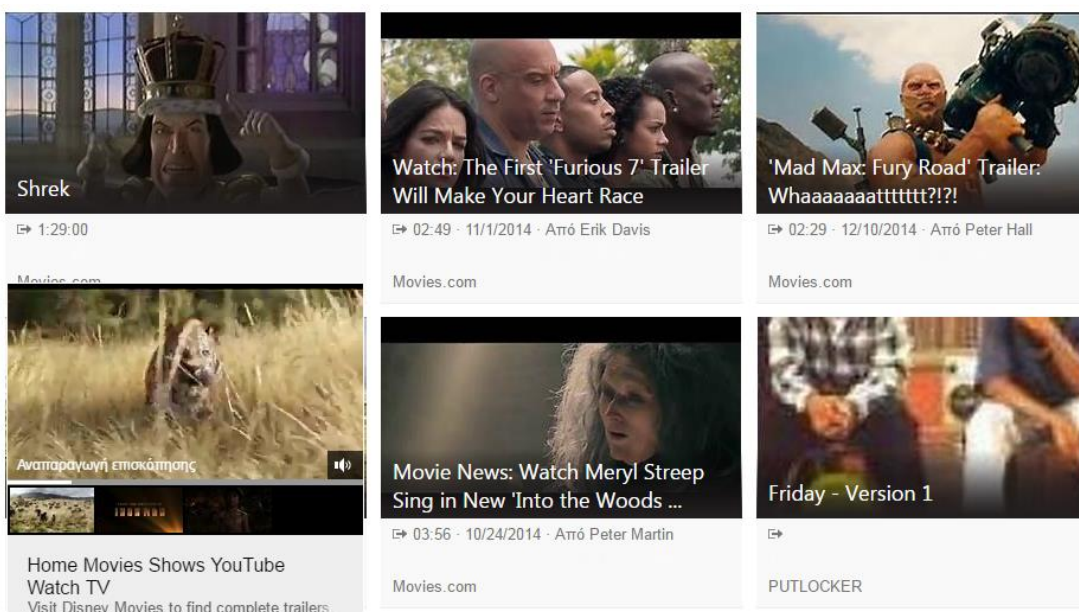
[Flights from New York, NY to Atlanta, GA](#)
[bing.com/travel](#)

From	To	Weekend Estimates
<input type="text" value="New York, NY (NYC)"/>	<input type="text" value="Atlanta, GA (ATL) -"/>	↑ \$239 Depart December 13
Leave	Return	↓ \$340 Depart December 20
<input type="text" value="01/03/2014"/>	<input type="text" value="01/05/2014"/>	↑ \$322 Depart December 27
<input type="button" value="Find flights"/>		

Εικόνα 2.5: Flight cost prediction της Bing[23].

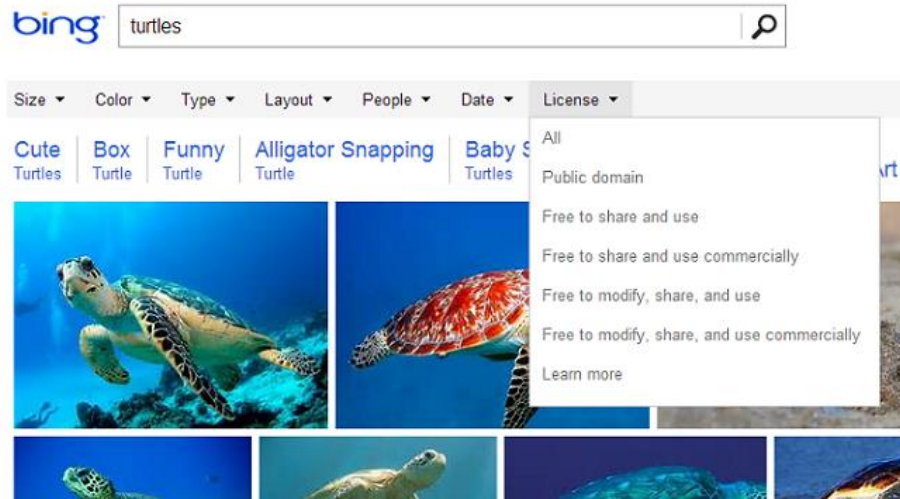
Unique Search Operators όπως LinkFromDomain: Το συγκεκριμένο χαρακτηριστικό της bing παρόλο που δεν έχει βελτιστοποιηθεί ώστε να εμφανίζεται σε όλες τις αναζητήσεις του χρήστη, μπορεί να αποδειχθεί πολύ χρήσιμο. Όπως είναι γνωστό μερικές MA δίνουν τη δυνατότητα αναζήτησης ψάχνοντας μόνο ένα συγκεκριμένο τύπο αρχείου(filetype). Ως αποτέλεσμα η MA θα μας επιστρέψει αποτελέσματα με βάση το τύπο του αρχείου που του έχουμε πληκτρολογήσει. Το linkfromdomain θα μας επιστρέψει τις πιο δημοφιλείς και αγαπημένες σελίδες του blog ή της σελίδες που θα του ορίσουμε.

Video search : Κανείς δε θα μπορούσε να αρνηθεί πως ο βασιλιάς των βίντεο είναι το youtube. Όμως η bing με την αναζήτηση βίντεο προσφέρει τη λειτουργία να μπορείς να βλέπεις ένα trailer του βίντεο καθώς κάνεις hover πάνω από αυτό.



Εικόνα 2.6: Feature video search της bing

Photo Licence: Άλλο ένα πολύ ενδιαφέρον της MA bing είναι ότι προσφέρει τη δυνατότητα αναζήτησης φωτογραφιών ανάλογα με τα πνευματικά δικαιώματά τους. Το συγκεκριμένο φίλτρο αποτελεί βασική λειτουργία της αναζήτησης εικόνων της bing.



Εικόνα 2.7: Photo license λειτουργία της bing[23].

2.3 M.A της Baidu

Η MA της baidu αποτελεί μία κινέζικη υπηρεσία με έδρα το Πεκίνο. Η ίδια ιδρύθηκε το 2000 από τους κινέζους Robin Li και Eric Xu και προσφέρει αρκετές υπηρεσίες όπως αναζήτηση σε κινέζικες ιστοσελίδες, αρχεία ήχου και εικόνες. Συνολικά έχει 57 υπηρεσίες μεταξύ των οποίων αξίζει να αναφέρουμε την Baidu Baike(εγκυκλοπαίδεια) καθώς και ένα keyword-based forum.

2.3.1 Ονομασία της Baidu

Το όνομα baidu προέρχεται από ένα κινέζικο ποίημα το οποίο χρονολογείται ότι γράφτηκε πριν από 800 χρόνια επί της δυναστείας Song. Πιο συγκεκριμένα προέρχεται από το κλασικό ποίημα "Green Jade Table in The Lantern Festival" του Xin Qiji's. Η φράση που ενέπνευσε τους δημιουργούς της στη σημερινή ονομασία είναι "Having searched thousands of times in the crowd, suddenly turning back, She is there in the dimmest candlelight". Όπου η κυριολεκτική σημασία του είναι: οι εκατοντάδες φορές, αντιπροσωπεύουν την επιμονή για αναζήτηση του ιδανικού .



Εικόνα 2.8: Λογότυπο της Baidu[25].

2.3.2 Λόγοι για τους οποίους η MA Baidu έχει ξεχωρίσει

Η MA της baidu έχει διαφορετική προσωπικότητα σε σχέση με τη MA της google. Αν η MA της google μας ρωτάει “what can i do for you” (τι μπορώ να κάνω για σένα) η baidu μας ρωτάει “what can i help you to do”(τι μπορώ να σε βοηθήσω για να κάνεις). Η κινέζικη baidu έχει διαφορετικό τρόπο προσέγγισης και προσαρμόζεται στη συμπεριφορά του χρήστη. Πιο συγκεκριμένα αν μπορούσαμε να κατονομάσουμε τους λόγους για τους οποίους διαφέρει η baidu είναι οι εξής :

- Ο τρόπος με τον οποίο κατατάσσονται οι ιστοσελίδες είναι διαφορετικός. Οι διαχειριστές των ιστοσελίδων εκτός της βελτιστοποίηση των ιστοσελίδων τους θα πρέπει να πληρώσουν και για διαφήμιση για να εμφανίζονται πρώτοι. Σε αντίθεση με άλλες MA οι οποίες εμφανίζουν στο χρήστη ανακατεμένα αποτελέσματα (οργανικά και μη βλ. κεφάλαιο 6.1).

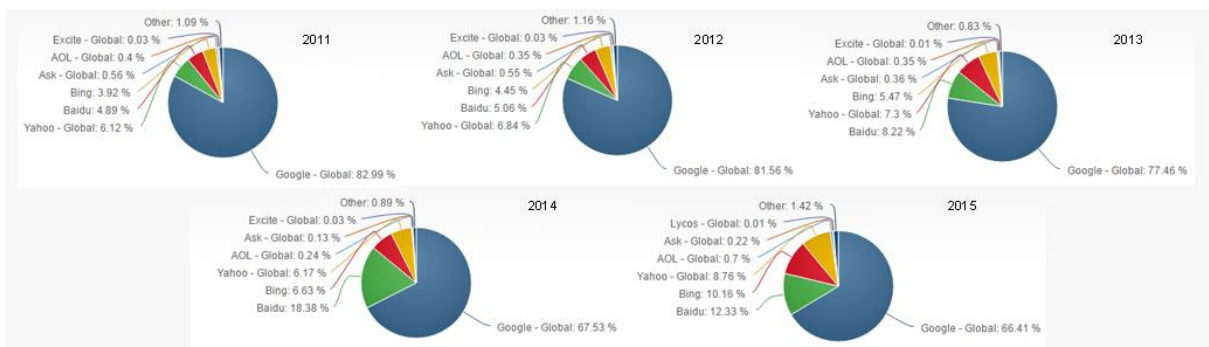
Η βελτιστοποίηση των ιστοσελίδων για τη MA καθορίζουν το ποσό το οποίο πρέπει να πληρώσει η ιστοσελίδα για να εμφανίζεται υψηλότερα στη κατάταξη.

- Η baidu είναι εστιασμένη στο πληθυσμό της Κίνας καθώς οι σελίδες οι οποίες αναδεικνύει στα υψηλά αποτελέσματα είναι κινέζικής γλώσσας και φιλοξενούνται(hosted) εντός της Κίνας.
- Οι σύνδεσμοι σε λίστες επί πληρωμή (paid link-building όπως θα δούμε παρακάτω) είναι διαθέσιμες αλλά δεν είναι τόσο δημοφιλής στους χρήστες.

- Δίνεται η δυνατότητα εκστρατείας (campaign) σε συγκεκριμένη περιοχή ώστε να προωθούνται online καταστήματα τα οποία μπορεί να μην έχουν φυσική υπόσταση και που ο ανταγωνισμός μπορεί να είναι πολύ λιγότερος.
- Παράγοντα κλειδί για τους οποίους διαφέρουν ανατολή και δύση αποτελεί η διαφήμιση. Καθώς οι εταιρείες καλούνται να πληρώσουν για να βρίσκονται πρώτες στη MA της baidu, ταυτόχρονα αποκτούν κύρος και εμπιστοσύνη. Ως αποτέλεσμα οι άνθρωποι που τη χρησιμοποιούν δεν είναι τόσο καχύποπτοι για τις αγορές μέσω αυτής. Για τις περισσότερες επιχειρήσεις η δημιουργία ενός καλού ονόματος και η γνωστοποίηση τους μέσα από το baidu και από τις υπόλοιπες MA αποτελεί διαφορετική προσέγγιση. Συμπερασματικά οι επιχειρήσεις χρειάζεται να επενδύσουν στο SEO για να φαίνονται υψηλά στις MA όπως της google ενώ για τη baidu αυτό δεν είναι αρκετό. Ο συνδυασμός SEO και πληρωμής είναι ο μοναδικός τρόπος ο οποίος μπορεί να εξασφαλίσει την αναγνωρισιμότητα που χρειάζονται οι εταιρείες[24].

2.4 Ποσοστά μεριδίου χρήσης που κατέχουν οι MA ανά τον κόσμο

Τα ποσοστά χρήσης των MA έχουν εξαχθεί από την ιστοσελίδα www.netmarketshare.com για τα τελευταία 5 χρόνια. Οι αναλύσεις είναι για όλες τις MA ανεξάρτητα του λειτουργικού συστήματος που χρησιμοποιήθηκε. Τα αποτελέσματα τα οποία εξάγει η σελίδα είναι για τις 7 πιο χρησιμοποιούμενες MA, ενώ κατατάσσει όλες τις υπόλοιπες στη κατηγορία other.



Εικόνα 2.9: Στατιστικά δεδομένα χρήσης μηχανών αναζήτησης 2011-2015.

Παρατηρώντας τα γραφήματα θα δούμε πως η google αποτελεί τη κυρίαρχη δύναμη μεταξύ των υπολοίπων ΜΑ. Όμως με το πέρασμα του χρόνου διαπιστώνουμε πως έχει φθίνουσα πορεία καθώς από την υπερισχύουσα θέση που βρισκόταν το 2011 με ποσοστό 83%, τη χρονιά 2015 κατέχει 66,4%. Αυτό που παρατηρούμε είναι πως και άλλες ΜΑ μπαίνουν δυναμικά στο χώρο με αποτέλεσμα να μειώσουν τη δυναμικότητα της google.

2.5 Λόγοι δημοφιλίας κάποιων Μ.Α που οδήγησαν στη σημερινή διαμόρφωση

Ο κυριότερος λόγος ο οποίος έχει καταστεί η Google η πιο δημοφιλής ΜΑ είναι διότι ήταν πολύ καλύτερη σε σχέση με τους ανταγωνιστές τις το 1998. Πολλοί άνθρωποι αποδίδουν την επιτυχία της google στην επαναστατική τεχνολογία που έφερε για τις ΜΑ. Πράγματι το κλειδί στην επιτυχία της, οφείλεται στο πείσμα που έδειξε η εταιρία καθώς αρνούσαν να ακολουθήσει τις ορθόδοξες τεχνικές της εποχής εκείνης. Όπως είδαμε η google αποτέλεσε καταλυτικό ρόλο στην εξέλιξη των ΜΑ καθώς οι τεχνικές της έχουν υιοθετηθεί και από άλλες ΜΑ. Πολλοί σημαντικό ρόλο στην επιτυχία των google, bing και baidu αποτελούν οι εξής λόγοι:

- Καλύτερη και ευκολότερη εμπειρία αναζήτησης: η ποιότητα των αποτελεσμάτων σε συνδυασμό με τη ταχύτητα με την οποία εξυπηρετούνται τα ερωτήματα δίνουν των χρηστών, έκαναν τις παραπάνω ΜΑ να διακριθούν σε σχέση με άλλες
- Μεγάλη προβολή των συγκεκριμένων ΜΑ: μέσα από τις δισεκατομμύρια αναζητήσεις που γίνονται διαφημίζονται οι υπηρεσίες και τα προϊόντα αυτών
- Καινοτόμα προϊόντα και υπηρεσίες: Εκτός της αναζήτησης προσφέρουν και άλλες ποικίλες web-based υπηρεσίες και προϊόντα. Τα πιο δημοφιλές είναι η υπηρεσία ηλεκτρονικού ταχυδρομείου, ιστοσελίδες κοινωνικής δικτύωσης οι χάρτες και άλλα. [20],[21].

3 Μέθοδοι, κανόνες και τεχνικές βελτιστοποίησης

Στο κεφάλαιο αυτό θα δούμε πως εξελίχθηκε το SEO και έφτασε στο σημείο που βρίσκεται σήμερα, τους “χρυσούς” κανόνες του, τις τεχνικές που χρειάζεται να εφαρμόσουμε για να βελτιώσουμε τη κατάταξη μιας ιστοσελίδας στα αποτελέσματα των ΜΑ(SERP) καθώς και τι πρέπει να αποφύγουμε. Αρχικά η λέξη SEO αποτελεί τη συντομογραφία των λέξεων Search Engine Optimization. Το SEO είναι μια διαδικασία η οποία επηρεάζει την προώθηση μιας ιστοσελίδας στα οργανικά αποτελέσματα(βλ. κεφάλαιο 6.1) των ΜΑ. Όσο πιο ψηλά εμφανίζεται μια ιστοσελίδα τόσο περισσότεροι θα είναι οι επισκέπτες της και οι δυνητικοί της πελάτες. Το SEO είναι μια διαφορετική διαδικασία από αυτή της διαφήμισης. Στη διαφήμιση πληρώνουμε για να εμφανιστεί η σελίδα μας ενώ στο SEO δεν πληρώνουμε, ο ιστοτοπός μας εμφανίζεται στα “φυσικά” αποτελέσματα των ΜΑ με τη σειρά που αποφασίζουν οι τελευταίες. Το SEO μπορεί να στοχεύει διαφορετικά είδη αναζήτησης όπως για εικόνες, videos, τοπικό περιεχόμενο, ακαδημαϊκά άρθρα, καθώς και αναζητήσεις ανά κλάδο.

3.1 Θεμελιώδης αλλαγές SEO με την έλευση των Panda και Penguin

Η καθοριστική ημερομηνία για το SEO είναι η 11η Φεβρουαρίου του 2011. Τότε η Google άρχισε να χρησιμοποιεί το Panda το οποίο είχε σχεδιαστεί για να φιλτράρει τις ιστοσελίδες χαμηλής ποιότητας. Αυτή η εξέλιξη άλλαξε τα δεδομένα στη βελτιστοποίηση των ιστοσελίδων. Μέχρι τότε αρκούσαν μερικά τρικ για να εμφανιστεί μια σελίδα ψηλά στα αποτελέσματα αναζήτησης. Με την υιοθέτηση του Panda τα πάντα άλλαξαν. Πολλοί διαχειριστές είδαν τους ιστοτόπους τους να τιμωρούνται από την Google και να εμφανίζονται πολύ χαμηλά στα αποτελέσματα αναζήτησης.

Panda

Ο Panda είναι μία επέκταση(extension) του αλγόριθμου της Google το οποίο επικεντρώνεται στην ποιότητα του περιεχομένου μιας ιστοσελίδας. Εμφανίστηκε για πρώτη φορά τον Φεβρουάριο του 2011 και από τότε έχουν γίνει αρκετές επικαιροποιήσεις του.



Εικόνα 3.1: Google panda[30].

Penguin

Τα πράγματα έγιναν ακόμα πιο δύσκολα τον Απρίλιο του 2012 όπου η Google άρχισε να χρησιμοποιεί το Penguin. Εάν το Panda ήταν μια σημαντική εξέλιξη το Penguin ήταν δέκα φορές πιο σημαντική. Άλλαξε εντελώς τον τρόπο με τον οποίο σκεφτόμαστε το Search Engine Optimization. Το Penguin έχει σχεδιαστεί για να εντοπίζει τις ιστοσελίδες στις οποίες έχουν χρησιμοποιηθεί τεχνικές βελτιστοποίησης που δεν είναι νόμιμες(black-hat seo)[30].



Εικόνα 3.2: Google penguin.

3.2 “Χρυσό κανόνες” SEO

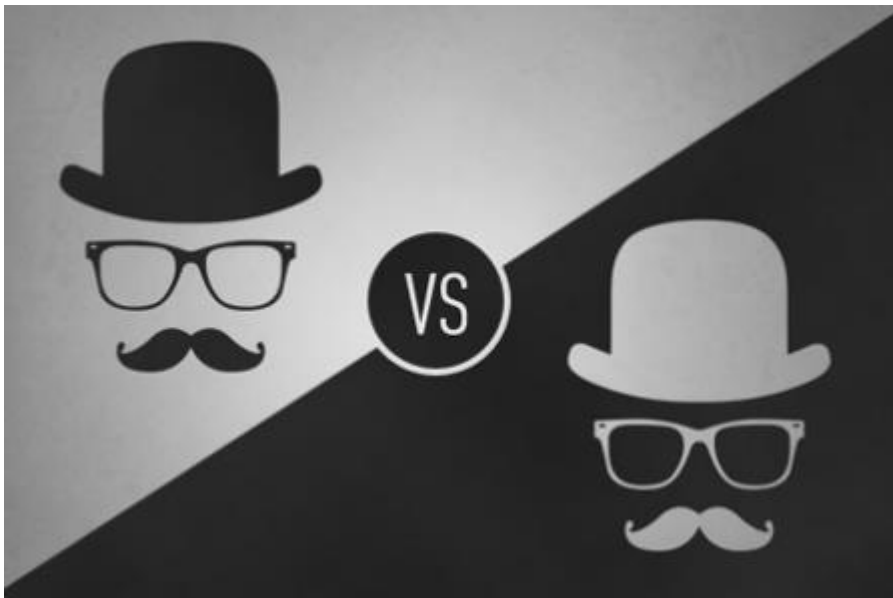
Η έλευση των google panda και google penguin διαμόρφωσε το seo όπως είναι σήμερα. Σύμφωνα με αυτά, θα μπορούσαμε να πούμε με βεβαιότητα πως πολύ βασικά αν όχι τα βασικότερα χαρακτηριστικά θα ήταν:

- Η ποσότητα των εισερχόμενων συνδέσμων (link) στην ιστοσελίδα μας: Αυτό πρακτικά σημαίνει πως όσο θα “μιλούσαν”-αναφερόντουσαν ξένες ιστοσελίδες, για την δική μας τόσο πιο αξιόλογοι και σημαντικοί θα ήμασταν για τους χρήστες του διαδικτύου.
- Η ποιότητα των εισερχόμενων συνδέσμων (link) στην ιστοσελίδα μας: Φυσικά, η ποσότητα των εισερχόμενων συνδέσμων, δεν είναι τόσο σημαντική αν αυτοί δεν προέρχονται από άλλους ιστότοπους που θεωρούνται και είναι αξιόλογοι προς τους χρήστες.
- Λιγότερο σημαντικό χαρακτηριστικό σε σχέση με τα δύο προηγούμενα αλλά αρκετά βασικό, είναι η ίδια η ιστοσελίδα. Ο τρόπος με τον οποίο θα προσπαθούν οι χρήστες να βρουν μια ιστοσελίδα μέσω συγκεκριμένων λέξεων κλειδιών (keywords) θα πρέπει να αντικατοπτρίζει την ίδια. Επίσης όπως οι χρήστες βλέπουν οπτικά μέσω διάφορων γραφικών μια ιστοσελίδα και θα τους αρέσει έτσι πρέπει να φαίνεται και στη ΜΑ. Αυτό θα συμβεί

ακολουθώντας τις προδιαγραφές του white-hat seo το οποίο θα αναλύσουμε παρακάτω στο ίδιο κεφάλαιο.

3.3 Μέθοδοι βελτιστοποίησης

Οι μέθοδοι βελτιστοποίησης ανάλογα με το αν χρησιμοποιούνται θεμιτοί ή αθέμιτοι τρόποι βελτιστοποίησης χωρίζονται σε white-hat seo και black-hat. Στη συγκεκριμένη ενότητα θα εξηγηθούν οι όροι αυτοί καθώς και τεχνικές που πρέπει να αποφεύγονται κατά τη βελτιστοποίηση μιας σελίδας.



Εικόνα 3.3: Black-hat και white-hat seo.

3.3.1 White hat SEO

Το White Hat SEO αφορά στις εγκεκριμένες στρατηγικές Search Engine Optimization που επιτρέπεται να χρησιμοποιήσουμε για να έχουμε την προώθηση της ιστοσελίδας μας σε υψηλότερες θέσεις στα αποτελέσματα αναζήτησης. Η Google δίνει σχετικές οδηγίες στους διαχειριστές(webmasters) για αυτές τις στρατηγικές, οι οποίες είναι οι ακόλουθες:

- Πρέπει να βεβαιωθούμε ότι η ιστοσελίδα έχει μια ξεκάθαρη δομή (information architecture) καθώς και συνδέσμους(links) με κείμενο. Κάθε σελίδα πρέπει να μπορεί να εντοπιστεί από ένα τουλάχιστον στατικό σύνδεσμο(link) με κείμενο. Η δομή αυτή δημιουργείται στο στάδιο της

- δημιουργίας του ιστότοπου και αποσκοπεί στην ικανοποίηση των αναγκών των επισκεπτών του.
- Ύπαρξη ενός sitemap για την ευκολότερη περιήγηση των χρηστών σε μία ιστοσελίδα.
 - Οι σύνδεσμοι(links) σε κάθε σελίδα σας θα πρέπει να είναι σε λογικό αριθμό, καλό είναι να αποφεύγονται οι υπερβολές.
 - Πρέπει να προσφέρουμε χρήσιμη και πλούσια πληροφόρηση και να δημιουργούμε σελίδες που με ακρίβεια και σαφήνεια περιγράφουν το αντικείμενο που πραγματεύονται.
 - Να λάβουμε υπόψιν τις λέξεις που χρησιμοποιούν οι χρήστες για να βρουν την ιστοσελίδα μας και να βεβαιωθούμε ότι αυτές περιλαμβάνονται στο κείμενο της. Είναι σημαντικό αυτά τα keywords να εμφανίζονται με φυσικό τρόπο στο κείμενο, γιατί τότε μόνο θεωρείται πρακτική White Hat SEO.
 - Να προτιμάμε το κείμενο από τις εικόνες για την απεικόνιση σημαντικής πληροφορίας. Οι αράχνες δεν μπορούν να αναγνωρίσουν το κείμενο το οποίο περιέχεται σε εικόνες. Εάν πρέπει να χρησιμοποιήσουμε εικόνες τότε πρέπει να γίνει χρήση του "ALT" χαρακτηριστικό(attribute) για να τη περιγραφή τους.
 - Τα <title> και ALT χαρακτηριστικά να είναι περιγραφικά και ακριβή.
 - Να ελέγξουμε για συνδέσμους οι οποίοι παρουσιάζουν πρόβλημα(broken links) και την σωστή σύνταξη της HTML.
 - Για τις εικόνες(images) που χρησιμοποιούμε πρέπει να δώσουμε περιγραφικό τίτλο, όπως my-thesis-image-example.jpg διότι περιέχει πιο πολύ πληροφορία από το IMG00001.JPG.
 - Η εκτεταμένη χρήση JavaScript, cookies, session IDs, frames, DHTML ή Flash μπορεί να δημιουργήσει προβλήματα στην ευρετηρίαση(indexing) της ιστοσελίδας μας.
 - Πρέπει να είμαστε σίγουροι ότι επιτρέπουμε στα Google bots να ανιχνεύσουν την ιστοσελίδα μας χωρίς session IDs ή arguments τα οποία καταγράφουν την πορεία τους στην ιστοσελίδα μας.
 - Ο web server μας υποστηρίζει το If-Modified-Since HTTP header. Αυτό το χαρακτηριστικό επιτρέπει στον web server μας να πει στην Google αν το περιεχόμενο της ιστοσελίδας έχει αλλάξει μετά την τελευταία επίσκεψη των αραχνών.

- Χρήση του robots.txt αρχείου για να την καθοδήγηση των αραχνών ποιες σελίδες να ανιχνεύσουν και ποιες όχι.
- Αν χρησιμοποιούμε διαφημίσεις στην ιστοσελίδα μας πρέπει να είμαστε σίγουροι ότι δεν μπλοκάρουν τις MA.
- Εάν χρησιμοποιούμε ένα σύστημα διαχείρισης περιεχομένου(content management system/cms) πρέπει να είμαστε σίγουροι ότι το σύστημα δημιουργεί σελίδες και συνδέσμους τους οποίους μπορούν να ανιχνεύσουν οι MA.
- Η ιστοσελίδα μας πρέπει να φαίνεται σωστά σε διαφορετικούς φυλλομετρητές(browsers).
- Να γίνει βελτιστοποίηση του χρόνου που χρειάζεται για να φορτωθεί η ιστοσελίδα μας.
- Ο σχεδιασμός των σελίδων μας πρέπει να είναι πρωτίστως για τους χρήστες και μετά για τις MA. Η βασική ερώτηση που πρέπει να κάνουμε είναι αν η οποιαδήποτε αλλαγή που σκεφτόμαστε θα είναι χρήσιμη και θα βοηθήσει την εμπειρία περιήγησης των χρηστών μας(user-experience/ UX)[34],[35].

3.3.2 Black hat SEO

Οι τεχνικές Black Hat SEO αποσκοπούν στην προώθηση ιστοσελίδων σε υψηλότερες θέσεις στα οργανικά αποτελέσματα(βλ. κεφάλαιο 6.1) των MA χρησιμοποιώντας μη αποδεκτές πρακτικές. Για το λόγο αυτό οι ιστοσελίδες που τις υιοθετούν τιμωρούνται από την Google και τις λοιπές MA. Πλέον έχουν δημιουργηθεί οι panda και penguin ώστε ο μόνος τρόπος για να αναδειχθεί μια σελίδα είναι να έχει περιεχόμενο το οποίο είναι μοναδικό και ενδιαφέρον για τους επισκέπτες της. Οι πιο βασικές από αυτές τις μη αποδεκτές πρακτικές seo είναι οι ακόλουθες:

Παραγέμιση περιεχομένου με τη λέξη κλειδί(Keyword stuffing)

Είναι η πιο συνηθισμένη τεχνική black hat seo. Αναφέρεται στην πρακτική της επανάληψης keywords σε μια προσπάθεια να ξεγελαστούν οι ΜΑ. Συχνά αυτά τα keywords εμφανίζονται σε λίστες(index) ή εμφανίζονται με αφύσικο τρόπο. Το να γεμίζουμε την ιστοσελίδα μας με keywords συνεπάγεται αρνητική εμπειρία για τους επισκέπτες της ενώ θα μειώσει την κατάταξη της στα οργανικά αποτελέσματα(βλ. κεφάλαιο 6.1). Για παράδειγμα:

Using **keyword** few times is good **keyword** use. Using **keyword** too many times is not good **keyword** use but bad **keyword** use. Using too many **keywords** is bad **keyword** use as It is **keyword** overuse, **keyword** kill or **keyword** stuffing. Bad **keyword** use, **keyword** repetition, **keyword** stuffing, **keyword** kill, or **keyword** overuse is bad **keyword** use.

Εικόνα 3.4: Παράδειγμα keyword stuffing.

Σελίδες με σκοπό να οδηγήσουν τους χρήστες σε άλλες(Doorway pages)

Είναι σελίδες φτωχού περιεχομένου που έχουν βελτιστοποιηθεί για συγκεκριμένες λέξεις κλειδιά και οι οποίες οδηγούν σε άλλη σελίδα από αυτή που αναζητούν οι χρήστες. Τείνουν να ενοχλούν τους χρήστες γιατί τους οδηγούν σε περιεχόμενο διαφορετικό από αυτό που ζήτησαν.

Κρυμμένα κείμενα(Hidden text)

Αφορά στην τοποθέτηση λιστών με keywords στο ίδιο χρώμα με το φόντο της ιστοσελίδας έτσι ώστε να μην είναι ορατές από τους χρήστες αλλά να ανιχνεύονται από τις ΜΑ. Είναι μια προσπάθεια να ξεγελαστούν οι τελευταίες και για αυτό τιμωρείται. Ίσως η πιο χρησιμοποιημένη τεχνική Black Hat SEO.

Συμμετοχή σε συνδέσμους λιστών(Participating in link schemes)

Αφορά σε σύνδεση(linking) με ιστοσελίδες των οποίων ο μοναδικός σκοπός τους είναι να προβάλλουν λίστες με συνδέσμους συχνά άσχετων μεταξύ

Τεχνικές, μέθοδοι και κανόνες για τη βελτιστοποίηση ανάπτυξης ιστοσελίδων

τους σελίδων. Δεν έχουν καμία επίδραση στη βελτίωση κατάταξης ενώ διατρέχουμε τον κίνδυνο να τιμωρηθούμε για αυτή την πρακτική.

Αυτόματη παραγωγή περιεχομένου(Automatically generated content)

Είναι η παραγωγή κειμένου με αυτοματοποιημένο τρόπο όπως είναι η δημιουργία παραγράφων με τυχαίο κείμενο το οποίο δεν έχει αξία για τον αναγνώστη του αλλά το οποίο περιέχει keywords. Κείμενο αυτής της κατηγορίας θα μπορούσε να δημιουργηθεί με διάφορες τεχνικές, όπως είναι μέσω RSS feeds, με αποτελέσματα αναζήτησης, με τη χρήση συνωνύμων, με αυτοματοποιημένες διαδικασίες όπως είναι τα Markov chains ή με αυτόματη μετάφραση η οποία δεν επιμελείται από ανθρώπους.

Παροχή διαφορετικού περιεχομένου στους χρήστες από αυτό που παρουσιάζεται στις MA (Cloacking)

Αφορά στην πρακτική κατά την οποία παρουσιάζεται διαφορετικό περιεχόμενο ή URLs στους χρήστες και διαφορετικό στις MA. Θεωρείται ότι παραβιάζει τις προδιαγραφές (Webmaster Guidelines) της Google[35],[44] γιατί παρέχει στους χρήστες διαφορετικά αποτελέσματα από αυτά που αναζητούν. Παράδειγμα cloacking είναι η παροχή σελίδας HTML text στις MA ενώ στους χρήστες παρουσιάζεται μια σελίδα με εικόνες ή flash[32].

3.4 Εφαρμογή των white-hat seo τεχνικών

Όσα έχουν ειπωθεί μέχρι τώρα θα πρέπει να εφαρμοστούν ξεχωριστά πάνω στη σελίδα(on site seo) αλλά και εκτός σελίδας (off site seo). Τα παραδείγματα τα οποία θα χρησιμοποιηθούν είναι από τις οδηγίες (guidelines) που έχει δημοσιεύσει η google.

3.4.1 On- site SEO

Οι on- site ή on-page τεχνικές αφορούν τις αλλαγές που πρέπει να γίνουν εντός της σελίδας, από το περιεχόμενο μέχρι και τους τίτλους, τις ετικέτες (tags), σύνδεσμο (url) και άλλα.

Δημιουργία μοναδικών και ακριβών τίτλων των σελίδων μας.

Ο τίτλος της σελίδας πρέπει να είναι όσο το δυνατό πιο περιγραφικός και να περιέχει λέξεις κλειδιά της σελίδας ώστε να ενημερώνει και τους χρήστες αλλά και την ίδια τη ΜΑ για το περιεχόμενο το οποίο πρόκειται να δουν. Για το λόγο αυτό θα πρέπει να είναι κατανοητός. Για κάθε μία σελίδα της ιστοσελίδας μας θα πρέπει να υπάρχει ο ξεχωριστός τίτλος που να υποδεικνύει το περιεχόμενο της. Η ετικέτα του τίτλου (<title>) στην html θα πρέπει να τοποθετείται μεταξύ των ετικετών head ως <title> ... </title>. Το μέγεθος του τίτλου είναι καλό να μην ξεπερνάει τους 50-60 χαρακτήρες. Στη παρακάτω εικόνα θα δούμε πως πρέπει να είναι ενσωματωμένος ο τίτλος της σελίδας μέσα σε ένα html έγγραφο:

```
<html>
<head>
<title>Brandon's Baseball Cards - Buy Cards, Baseball News, Card Prices</title>
<meta name="description" content="Brandon's Baseball Cards provides a
large selection of vintage and modern baseball cards for sale. We also offer
daily baseball news and events in">
</head>
<body>
```

Εικόνα 3.5: Μορφή του τίτλου μέσα στην HTML[35].

Χρησιμοποίηση της περιγραφής που προσφέρει η ετικέτα μετα (meta description).

Η συγκεκριμένη ετικέτα προσφέρει σε πολλές ΜΑ μια συνοπτική περιγραφή για το περιεχόμενο της σελίδας. Το μέγεθος της συγκεκριμένης περιγραφής είναι μεγαλύτερο από αυτό του τίτλου. Το ιδανικό μέγεθος είναι στους 155-160 χαρακτήρες μαζί με τα κενά. Για κάθε μια σελίδα είναι ωφέλιμο να υπάρχει

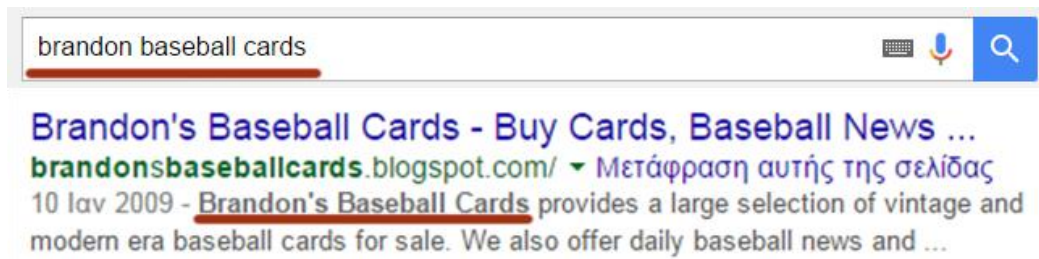
Τεχνικές, μέθοδοι και κανόνες για τη βελτιστοποίηση ανάπτυξης ιστοσελίδων

διαφορετική μετα-περιγραφή (meta description). Όπως η ετικέτα του τίτλου έτσι και η μετα-ετικέτα τοποθετείται μεταξύ της ετικέτας head.

```
<html>
<head>
<title>Brandon's Baseball Cards - Buy Cards, Baseball News, Card Prices</title>
<meta name="description" content="Brandon's Baseball Cards provides a
large selection of vintage and modern baseball cards for sale. We also offer
daily baseball news and events in">
</head>
<body>
```

Εικόνα 3.6: Meta description στην HTML[35].

Πρέπει να αποφεύγεται η γενίκευση στη περιγραφή. Σύμφωνα με το παράδειγμα μία περιγραφή όπως “Μια σελίδα με κάρτες του baseball”(Page about baseball cards) θεωρείται πολύ κακή. Επίσης η περιγραφή δε πρέπει να γεμίζεται μόνο με λέξεις κλειδιά όπως για παράδειγμα “κάρτες,baseball(cards,baseball) κτλ”. Τέλος η μετα-περιγραφή εμφανίζεται κατά την αναζήτηση στη MA και οι λέξεις που χρησιμοποιούνται κατά για την αναζήτηση είναι με έντονα γράμματα.



Εικόνα 3.7: Εμφάνιση μετα-ετικέτας και έντονα γράμματα αυτής.

Βελτίωση της δομής του συνδέσμου(URL) της σελίδας μας.

Η δημιουργία κατηγοριών και η σωστή ονοματοποίηση αυτών δε βοηθάει μόνο τον ίδιο το δημιουργό της σελίδας αλλά συμβάλλει και στη καλύτερη οργάνωση της σελίδας. Επιπλέον βοηθάει τις ίδιες της MA κατ της σάρωση(crawling) των αρχείων της ιστοσελίδας. Επίσης βοηθάει και τους ίδιους τους χρήστες καθώς ένας σύνδεσμος όπως στο πάνω μέρος της εικόνας 3.8 θα ήταν δύσκολο να τον θυμούνται.



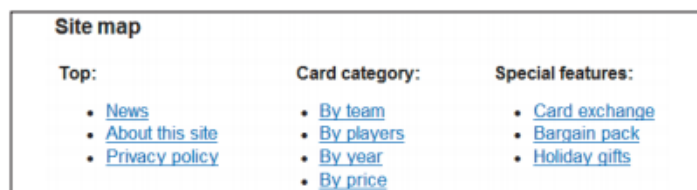
Εικόνα 3.8: Κακός σύνδεσμος url[35].



Εικόνα 3.9: Σύνδεσμος φιλικός προς τις MA.

Δημιουργία sitemap για την ευκολότερη περιήγηση της σελίδας.

Το sitemap θα μπορούσαμε να πούμε πως αποτελεί ένα πλήρη χάρτη της ιστοσελίδας μας. Η ύπαρξη του βοηθάει και τους χρήστες αλλά και τους σαρωτές(crawlers). Για το λόγο αυτό θα πρέπει να σχεδιάζονται 2 sitemaps. Όπως είναι λογικό ένας χάρτης της ιστοσελίδας βοηθάει στη περιήγηση των χρηστών, ιδιαίτερα όταν αυτοί δε βρίσκουν το περιεχόμενο το οποίο αναζητούν. Η δημιουργία του γίνεται με HTML για αυτό ονομάζεται και HTML sitemap. Σε αντίθεση με το παραπάνω χάρτη για τις MA σχεδιάζεται ένας XML sitemap.



Εικόνα 3.9: HTML sitemap για τη διευκόλυνση των χρηστών[35].

```
<?xml version="1.0" encoding="UTF-8"?>
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
  <url>
    <loc>http://www.brandonsbaseballcards.com/</loc>
    <changefreq>daily</changefreq>
    <priority>0.8</priority>
  </url>
  <url>
    <loc>http://www.brandonsbaseballcards.com/news/</loc>
  </url>
  <url>
    <loc>http://www.brandonsbaseballcards.com/news/2008/</loc>
  </url>
  <url>
    <loc>http://www.brandonsbaseballcards.com/news/2009/</loc>
  </url>
  <url>
    <loc>http://www.brandonsbaseballcards.com/news/2010/</loc>
  </url>
</urlset>
```

Εικόνα 3.11: XML sitemap για τη διευκόλυνση των ΜΑ[35].

Προσφορά πλούσιου και πρωτότυπου περιεχομένου στις διάφορες σελίδες χρησιμοποιώντας τις λέξεις κλειδιά.

Εδώ μιλάμε για καθαρό κείμενο χωρίς να υπολογίζουμε τον κώδικα html. Εύλογα κάποιος θα αναρωτηθεί αν αυτός ο μεγάλος αριθμός λέξεων μπορεί να μειώσει την εμπειρία του χρήστη (user experience-ux) γιατί η σελίδα που θα το περιέχει θα είναι μεγάλη σε μέγεθος και ο τελευταίος θα χρειαστεί να κάνει αρκετό scroll down για να το διαβάσει. Το θέμα όμως αυτό μπορεί να αντιμετωπιστεί χρησιμοποιώντας καρτέλες (tab navigation) στη συγκεκριμένη σελίδα. Με τον τρόπο αυτό και όλο το κείμενο είναι σε μία σελίδα και η ανάγνωση του θα είναι εύχρηστη. Δε θα πρέπει να μείνουμε μόνο στο μέγεθος του κειμένου αλλά και στη πρωτοτυπία του.

Η Google εξετάζει αν το περιεχόμενο μιας σελίδας επαναλαμβάνεται είτε στην ίδια ιστοσελίδα είτε σε άλλες. Υπάρχουν διάφορες τεχνικές για αυτό όπως είναι το hashing ή η καταγραφή του μοτίβου των λέξεων. Σε κάποιες περιπτώσεις, όπως στα ηλεκτρονικά καταστήματα (e-shops) αυτό είναι ένα πραγματικό

πρόβλημα γιατί για παράδειγμα στις περιγραφές των προϊόντων χρησιμοποιείται κείμενο το οποίο υπάρχει συνήθως στην ιστοσελίδα του κατασκευαστή και το οποίο στις περισσότερες περιπτώσεις είναι το ίδιο και σε άλλα ηλεκτρονικά καταστήματα που έχουν τα ίδια προϊόντα. Αυτό μπορεί να αντιμετωπιστεί με τον εμπλουτισμό του κειμένου με αξιολογήσεις που κάνουν οι πελάτες, με την αλλαγή του αρχικού κειμένου και αν είναι δυνατόν και των φωτογραφιών. Σε τελευταία ανάλυση αυτό που δεν θέλουμε σε καμία περίπτωση είναι το κείμενο της σελίδας να υπάρχει το ίδιο και απaráλλαχτο σε πολλές άλλες ιστοσελίδες. Εάν τιμωρηθούμε για αντιγραφή τότε το να το ξαναγράψουμε μπορεί να μην είναι αρκετό για να αρθεί η τιμωρία. Όπως έχει δηλώσει ο Matt Cutts, επικεφαλής του webspam της Google, το νέο κείμενο θα πρέπει να είναι υψηλής ποιότητας και να προσθέτει αξία στο διαδίκτυο. Μια μορφή διασφάλισης του πρωτότυπου περιεχομένου μας είναι να το συνδέσουμε με το Google+ profile μας (rel="author") ή με το Google+ page μας (rel="publisher"). Η διαδικασία αυτή ονομάζεται Google Authorship. Με τον τρόπο αυτό δείχνουμε στην Google ότι εμείς είμαστε οι δημιουργοί αυτού του πρωτότυπου περιεχομένου. Παρόλο που η συγκεκριμένη εταιρία έχει δηλώσει στις 28 Αυγούστου 2014 ότι δεν θα επεξεργάζεται πλέον το Google Authorship ταυτόχρονα, δεν μας αποτρέπει από το να συνεχίσουμε να το αξιοποιούμε. Είναι επίσης χρήσιμο να ελέγχετε ανά τακτά χρονικά διαστήματα αν κάποια άλλη ιστοσελίδα σας έχει αντιγράψει περιεχόμενο χωρίς να κάνει καμία αναφορά στην δική σας. Υπάρχουν διάφορα εργαλεία για το έργο αυτό όπως το copyscape. Αν εντοπιστεί ιστοσελίδα η οποία έχει αντιγράψει κάτι το οποίο είναι πρωτότυπο από εμάς, αυτό που πρέπει να κάνουμε είναι να επικοινωνήσουμε με το διαχειριστή της ώστε είτε να βάλει κάποιο σύνδεσμο που να κάνει ανακατεύθυνση στη πηγή(εμάς) είτε να αφαιρέσει το περιεχόμενο μας[30].

Βελτιστοποίηση των φωτογραφιών.

Η ύπαρξη εικόνων αποτελεί έναν απλό στοιχείο στις ιστοσελίδες, επιδέχεται όμως βελτιστοποιήσεις. Το όνομα των φωτογραφιών πρέπει να είναι μικρό και περιγραφικό. Τα ονόματα του τύπου "image1.jpg" πρέπει να αποφεύγονται και να μην είναι μεγάλα σε μέγεθος. Παράλληλα θα πρέπει να υπάρχει μια συνοπτική περιγραφή της εικόνας χρησιμοποιώντας το χαρακτηριστικό(attribute) alt. Οι συνηθέστερες μορφές εικόνων που συναντούμε είναι gif, bmp, jpeg και png με τις 2

Τεχνικές, μέθοδοι και κανόνες για τη βελτιστοποίηση ανάπτυξης ιστοσελίδων

τελευταίες από αυτές να είναι οι πιο δημοφιλείς. Επίσης καλό είναι να υπάρχει ένα sitemap εικόνων ώστε να παρέχονται περισσότερες πληροφορίες για το googlebot. Η δομή του sitemap είναι παρόμοια με αυτή της ιστοσελίδας σε XML. Η σωστή μορφή που μέσα στην html είναι :

```

```

Εικόνα 3.12: HTML εικόνα με alt text.

Βελτιστοποίηση των ετικετών επικεφαλίδων.

Οι επικεφαλίδες ορίζονται από <h1> μέχρι <h6> όπου ο αριθμός ορίζει το μέγεθος της γραμματοσειράς του κειμένου. Ο <h1> αποτελεί την ετικέτα με τη μεγαλύτερη σε μέγεθος γραμματοσειρά ενώ ο <h6> με τη μικρότερη. Αυτές χρησιμοποιούνται για τη καλύτερη παρουσίαση της σελίδας στο χρήστη. Επίσης δημιουργούν μια ιεραρχική δομή στο περιεχόμενο μας καθιστώντας το ευκολότερο για την πλοήγηση των χρηστών. Καλό είναι να μη χρησιμοποιείται αν δε χρειάζεται και η χρήση τους να μην είναι υπερβολική στο ίδιο κείμενο. Αντί αυτού αν θέλουμε να δώσουμε έμφαση σε κάποιο σημείο του κειμένου είναι συνετό να χρησιμοποιήσουμε τις ετικέτες ή .

```
</head>
<body>
<h1>Brandon's Baseball Cards</h1>
<h2>News - Treasure Trove of Baseball Cards Found in Old Barn</h2>
<p>A man who recently purchased a farm house was pleasantly surprised ...
dollars worth of vintage baseball cards in the barn. The cards were ... in news
papers and were thought to be in near-mint condition. After ... the cards to his
grandson instead of selling them.</p>
```

Εικόνα 3.13: Ορθή χρήση των ετικετών επικεφαλίδων[35].

Αποτελεσματική χρήση του αρχείου robots.txt

Το συγκεκριμένο αρχείο είναι πολύ σημαντικό διότι ενημερώνει τη MA για τις περιοχές της σελίδας που πρέπει να σαρώσει και να ευρετηριάσει. Αποτελεί ένα απλό αρχείο txt. Για την απαγόρευση ή αποτροπή σάρωσης απο τον crawler

χρησιμοποιείται η ετικέτα disallow. Το συγκεκριμένο αρχείο πρέπει να βρίσκεται στη ρίζα (root) της σελίδας μας (πχ www.myexample.com/robots.txt). Για τη προστασία ευαίσθητου ή απόρρητου περιεχομένου δεν ενδείκνυται η χρησιμοποίηση του αρχείου robots. Ένας λόγος είναι ότι άλλου είδους MA οι οποίες ονομάζονται rogue, δεν αναγνωρίζουν τις προδιαγραφές του αρχείου robots και συνεπώς παραβιάζουν τους κανόνες που αυτό θέτει. Μία καλή λύση προστασίας του περιεχομένου αποτελεί η κωδικοποίηση μέσω του αρχείου .htaccess. Η google προτείνει να αποφεύγονται να σαρωθούν οι σύνδεσμοι URL οι οποίοι δημιουργούνται ως αποτελέσματα των υπηρεσιών proxy[35].

```
# robots.txt
User-agent: *
Disallow: /*?
Disallow: /wp-
Disallow: /feed/
Disallow: */feed/
Disallow: /cgi-bin/
Disallow: /comments/
Disallow: */comments/
Disallow: /trackback/
Disallow: */trackback/
Disallow: /wp-admin/
Disallow: /wp-content/
Disallow: /wp-includes/
Disallow: /wp-login.php
Disallow: /wp-content/cache/
Disallow: /wp-content/themes/
Disallow: /wp-content/plugins/
Allow: /
Sitemap: http://msrtm.technerdia.com/sitemaps/msrtm-technerdia.xml
```

Εικόνα 3.14: Παράδειγμα αρχείου robots.txt.

Ρύθμιση της παραμέτρου rel="nofollow"

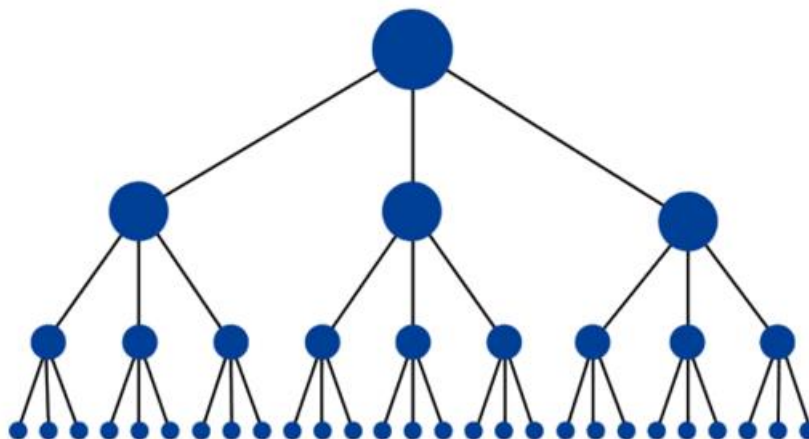
Η παράμετρος rel είναι συντομογραφίας της λέξης relationship. Η ρύθμιση της με παράμετρο "nofollow" ενημερώνει τη MA να μην ακολουθήσει οποιοδήποτε εξερχόμενο σύνδεσμο της σελίδας. Σύμφωνα με την ίδια τη google η συγκεκριμένη ρύθμιση πρέπει να χρησιμοποιείται σε περιπτώσεις όπως: για περιεχόμενο των σελίδων των οποίων δε μπορείτε να εγγυηθείτε. Αυτό βοηθάει στην αποφυγή ευρετηρίασης spam σελίδων οι οποίες θα έχουν αντίκτυπο και στη κατάταξη της ιστοσελίδας.


```
<html>
<head>
<title>Brandon's Baseball Cards - Buy Cards, Baseball News, Card Prices</title>
<meta name="description" content="Brandon's Baseball Cards provides a
large selection of vintage and modern baseball cards for sale. We also offer
daily baseball news and events in">
<meta name="robots" content="nofollow">
</head>
<body>
```

Εικόνα 3.15: Παράδειγμα ενσωμάτωσης της παραμέτρου στην HTML[35].

Εσωτερικοί σύνδεσμοι με ανακατευθύνσεις στη δική μας σελίδα

Οι εσωτερικοί σύνδεσμοι αποτελούν σημαντικό παράγοντα στη προώθηση και ανάδειξη της ιστοσελίδας. Αρχικά ως εσωτερικούς συνδέσμους ορίζουμε τους συνδέσμους οι οποίοι δείχνουν σε άλλες σελίδες της ίδιας ιστοσελίδας. Οι κυριότεροι λόγοι που χρησιμοποιούνται είναι διότι είναι γίνεται ευκολότερη η αναζήτηση αλλά και η σάρωση από το χρήστη και τον crawler αντίστοιχα και βοηθούν στην δόμηση της πληροφορίας. Η ιδανική δομή μιας ιστοσελίδας πρέπει να μοιάζει με μια πυραμίδα όπου η κορυφή είναι η αρχική σελίδα της ιστοσελίδας(homepage). Η συγκεκριμένη δομή επιτρέπει την ομαλή σάρωση ολόκληρης της ιστοσελίδας από τον crawler και αυξάνει τη πιθανότητα καλύτερης κατάταξης των σελίδων. Η συγκεκριμένη δομή είναι κοινή μεταξύ ιστοσελίδων με υψηλές κατατάξεις(πχ amazon) όπου συναντούμε ένα σύστημα από κατηγορίες και υποκατηγορίες αυτών[35],[36].



Εικόνα 3.16: Πρότυπη δομή ιστοσελίδας[36].

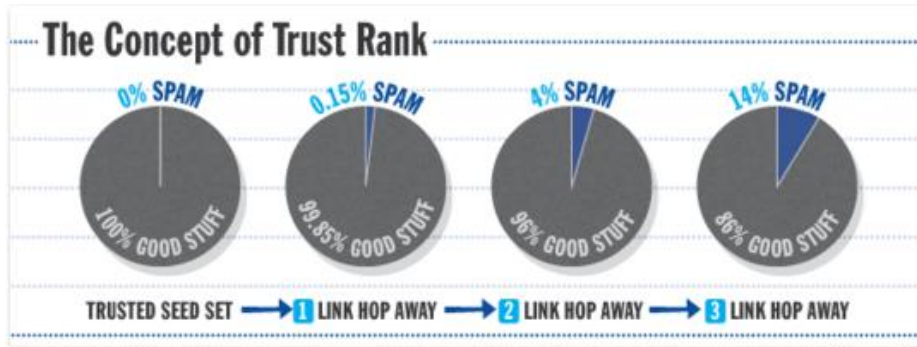
3.4.2 Off- site SEO

Σε αντίθεση με το on site seo το off site στοχεύει σε ενέργειες που πρέπει να γίνουν πέρα από την ιστοσελίδα μας. Το off-site είναι εξίσου σημαντικό με το on-site και η επιτυχία ανάδειξης εξαρτάται και από τα δύο. Οι ενέργειες αυτές είναι:

Χτίσιμο συνδέσμων προς την ιστοσελίδα μας(Link Building)

Σημαντική παράμετρος για την προώθηση των ιστοσελίδων στις MA αποτελεί η δημιουργία ποιοτικών backlinks. Η προώθηση ιστοσελίδων με το χτίσιμο συνδέσμων(links) είναι μια επίμονη και συνεχής διαδικασία. Πρόκειται για συνδέσμους που υπάρχουν σε άλλες ιστοσελίδες στους οποίους κάνοντας κλικ οδηγούμαστε στον ιστότοπο μας. Η ύπαρξη τους ερμηνεύεται από τις MA ως ψήφος εμπιστοσύνης της διαδικτυακής κοινότητας στη σελίδα μας. Για τη προώθηση ιστοσελίδων με ποιοτικά backlinks θα πρέπει αυτά να προέρχονται με τους εξής τρόπους:

- Από σελίδες με υψηλό PageRank το οποίο με απλά λόγια είναι η βαθμολογία που δίνει η Google για κάθε ιστοσελίδα, όσο υψηλότερη είναι τόσο πιο σημαντικό είναι το backlink από αυτή. Η βαθμολία αυτή κυμαίνεται από το 0-10. Το μειονέκτημα του PageRank είναι ότι έχει αρκετό καιρό να επικαιροποιηθεί από την Google, συνεπώς νέες αξιόλογες ιστοσελίδες έχουν PR 0. Για να αντισταθμίσουμε αυτό το μειονέκτημα του PR μπορούμε να χρησιμοποιήσουμε το TrustRank.
- Βρίσκονται σε ιστοσελίδες με υψηλό TrustRank. Για να αντιληφθούμε το σκεπτικό του υπολογισμού του ας δούμε την ακόλουθη εικόνα:



Εικόνα 3.17: Κατάταξη σελίδων με βάση τον αριθμό spam(Trustrank).

Στα αριστερά της εικόνας 3.17 βρίσκονται ιστοσελίδες με το μεγαλύτερο βαθμό εμπιστοσύνης, οι οποίες θεωρούνται ιδιαίτερα ποιοτικές και αξιόπιστες, όπως είναι εκείνη του BBC, New York Times, CNN κτλ. Αυτές θεωρούνται ότι έχουν 0% spam. Όσο πιο πολύ απέχει μια ιστοσελίδα από αυτές που θεωρούνται απόλυτα αξιόπιστες τόσο χαμηλότερο TrustRank θα έχει.

- Φιλοξενούνται σε σελίδες με περιεχόμενο αντίστοιχο με το δικό μας. Εάν το περιεχόμενο μας αφορά σε θέματα software και hardware τότε το να έχουμε backlink από ένα website που ασχολείται με λογιστικές υπηρεσίες φαίνεται περίεργο. Η αξία ενός τέτοιου συνδέσμου θεωρείται πλέον χαμηλή και σε μερικές περιπτώσεις μπορεί να αξιολογηθεί αρνητικά από τις MA.
- Έχουν διασπορά σε επίπεδο domain, δηλαδή προέρχονται από πολλές ιστοσελίδες και όχι από λίγες. Όσο μεγαλύτερη είναι η πρώτη τόσο αποτελεσματικότερη είναι η προώθηση ιστοσελίδων. Είναι προτιμότερο να έχουμε 10 συνδέσμους προς την σελίδα μας από 10 διαφορετικά websites παρά 100 backlinks από 3 ιστοτόπους.
- Έχουν διασπορά αναφορικά με τον τύπο των ιστοσελίδων από τα οποία προέρχονται. Για παράδειγμα αν όλα τα backlinks μας προέρχονται από καταλόγους(directories) τότε αυτά υποβαθμίζονται. Είναι λοιπόν σημαντικό τα backlinks μας να προέρχονται από διάφορους τύπους ιστοσελίδων. Προέρχονται κυρίως από TLDs (Top Level Domains) τα οποία αφορούν στην χώρα που στοχεύει το website. Για παράδειγμα η επιτυχημένη προώθηση ιστοσελίδων στην Ελλάδα προϋποθέτει backlinks από TLD σε .gr.

Δημιουργία ποιοτικών backlinks

Υπάρχουν δεκάδες τρόποι για την δημιουργία ποιοτικών backlinks, θα αναφέρουμε τους σημαντικότερους από αυτούς. Η λογική που προτείνουμε για την δημιουργία τους είναι το χτίσιμο backlinks τα οποία μπορούν να αυξήσουν απευθείας την επισκεψιμότητα στην ιστοσελίδα σας, όχι μόνο να επηρεάσουν θετικά την κατάταξη της.

Directories: Τα directories είναι κατάλογοι που μπορούμε να υποβάλλουμε την ιστοσελίδα μας. Η τελευταία θα ελεγχθεί από ανθρώπους και στην περίπτωση όπου πληρεί τα κριτήρια του καταλόγου θα δημοσιευθεί. Είναι σημαντικό να επιλέξουμε την πιο σχετική κατηγορία με την θεματολογία του ιστότοπου μας και τα στοιχεία επικοινωνίας – όνομα, διεύθυνση, τηλέφωνο να είναι συστηματικά τα ίδια. Επίσης πρέπει να αλλάζει η περιγραφή έτσι ώστε σε κάθε κατάλογο να είναι ελαφρώς διαφοροποιημένη. Η προώθηση ιστοσελίδων με εγγραφή σε ηλεκτρονικούς καταλόγους χρειάζεται προσοχή και τα χαμηλής ποιότητας directories πρέπει να αποφεύγονται.

Guest blogging: Αφορά στην υποβολή σχολίων σε blogs. Για να είναι επιτυχημένη η συγκεκριμένη προσέγγιση θα πρέπει πρώτα να εντοπίζουμε blogs τα οποία πληρούν δύο προϋποθέσεις: πρώτον να έχουν θεματολογία σχετική με εκείνη της ιστοσελίδας μας. Δεύτερον να στοχεύουν την ίδια γεωγραφική περιοχή. Το σχόλιο που θα υποβάλλουμε θα πρέπει να προσθέτει αξία στο post του blog. Αυτό σημαίνει ότι θα πρέπει να έχουμε διαβάσει το περιεχόμενο του και να γράψουμε κάτι σχετικό και εποικοδομητικό. Το κλειδί στο guest blogging είναι στη θέση του ονόματος αυτού που υποβάλλει το σχόλιο να βάζουμε λέξεις- κλειδιά για τις οποίες θέλουμε να επιτύχουμε υψηλή κατάταξη. Προσοχή αυτά τα keywords να διαφοροποιούνται, αν επαναλαμβάνουμε συνεχώς τα ίδια η Google θα μας τιμωρήσει. Η προώθηση ιστοσελίδων με το guest blogging δυσκολεύει πάντως διαρκώς δεδομένου ότι αρκετά blogs απενεργοποιούν τα σχόλια ή χρησιμοποιούν third-party commending systems όπως είναι το Disqus.

Δημοσιοποίηση του περιεχομένου μας: Όσοι περισσότεροι διαβάσουν το περιεχόμενο μας τόσο αυξάνονται οι πιθανότητες να δημιουργηθούν backlinks προς αυτό. Ένας τρόπος είναι η ενημέρωση με email(newsletters).

Social bookmarking: Υπάρχουν πολλές ιστοσελίδες που παρέχουν υπηρεσίες social bookmarking στην Ελλάδα και το εξωτερικό (πχ reddit). Η προώθηση ιστοσελίδων που επιτυγχάνεται πάντως με αυτή την πρακτική έχει χαμηλή δυναμική.

Δελτία τύπου: Αν έχουμε να αναδείξουμε κάτι το οποίο μπορεί να ενδιαφέρει αρκετό κόσμο τότε μπορούμε να γράψουμε ένα δελτίο τύπου για αυτό και βεβαίως να συμπεριλάβουμε σε αυτό το url της ιστοσελίδα σας. Όση περισσότερη δημοσιότητα επιτευχθεί τόσο περισσότεροι συνδέσμοι (links) θα δημιουργηθούν προς την σελίδα μας.

Wikipedia: Διερεύνηση αν υπάρχει κάποιο άρθρο στην Wikipedia σχετικό με το κλάδο μας. Αν όχι είναι μια χρυσή ευκαιρία να γράψουμε εμείς ένα. Έπειτα το δημοσιεύουμε και δημιουργούμε ένα backlink προς τον ιστότοπο μας το οποίο θα παραπέμπει σε περιεχόμενο το οποίο ενδιαφέρει τους αναγνώστες του. Η συγκεκριμένη τεχνική για προώθηση ιστοσελίδων θα πρέπει να χρησιμοποιείται με προσοχή, αποφεύγουμε την εμπορική προσέγγιση και δημιουργούμε περιεχόμενο που προσθέτει αξία.

Συγγραφή άρθρων: Απαιτητική εργασία δεδομένου ότι αναφερόμαστε σε πρωτότυπα άρθρα, τουλάχιστον 500 λέξεων. Μπορούν να τα υποβληθούν σε ιστοσελίδες που τα φιλοξενούν και αν δημοσιευθούν θα έχουμε χίτσει dofollow backlinks. Προσοχή τα άρθρα που υποβάλλουμε να μην έχουν δημοσιευθεί σε άλλες ιστοσελίδες γιατί τότε θα υπάρχει κίνδυνος τιμωρίας για αντιγραφή περιεχομένου[37].

Δημιουργία ενός αξιόπιστου Blog: Η δημιουργία και λειτουργία ενός blog παράλληλα με μια κεντρική ιστοσελίδα αποτελεί ένα πολύ καλό τρόπο για την αύξηση της αναγνωρισιμότητας μας. Επίσης η ύπαρξη ενός blog αποτελεί ένα πολύ καλό τρόπο δημιουργίας backlink προς τη κεντρική μας σελίδα, αυξάνοντας ταυτόχρονα και την επισκεψιμότητα της.

Μέσα κοινωνικής δικτύωσης(social media)

Οι ΜΑ αξιολογούν πλέον μια ιστοσελίδα και από την παρουσία της στα social media. Ανάλογα με το ύψος της δημοσιότητας που αποκτά μια ιστοσελίδα στα κοινωνικά δίκτυα παρατηρείται αλλαγή και στη κατάταξη της. Συνεπώς οι ενέργειες που πρέπει να κάνουμε είναι:

- Δημιουργία σελίδων στα μέσα κοινωνική δικτύωσης όπως, facebook, twitter, linkedin, google+ και youtube. Όπως καταλαβαίνετε όσα περισσότερα τόσο καλύτερα.
- Προσθήκη κουμπιών στην ιστοσελίδα μας ώστε να επιτρέπεται εύκολα στους χρήστες η κοινοποίηση.
- Προσθήκη like κουμπιών στη σελίδα.
- Βελτιστοποίηση της σελίδας για όλες τις κινητές συσκευές(smartphone, tablets κτλ).

Για να έχουμε όσο το δυνατό περισσότερη κίνηση από τα μέσα κοινωνικής δικτύωσης προς τη σελίδα μας, θα πρέπει να προκαλούμε το χρήστη ώστε να κάνει κλικ (click bait) για να διαβάσει το περιεχόμενό μας. Αν αυτό γίνει με επιτυχία τότε οι χρήστες ανταποκρίνονται με έναν ή με περισσότερους από του παρακάτω τρόπους:

- Αυξάνεται ο αριθμός των ανθρώπων που κάνουν like.
- Αυξάνεται ο αριθμός των ανθρώπων που κάνουν Facebook shares.
- Αυξάνεται ο αριθμός των ανθρώπων που κάνουν σχολίων στο Facebook.
- Αυξάνεται ο αριθμός των ανθρώπων που κάνουν των followers στο Twitter.
- Αυξάνεται ο αριθμός των tweets που αναφέρουν την ιστοσελίδα μας ή που περιλαμβάνουν ένα link προς αυτή
- Αυξάνεται ο αριθμός των ανθρώπων που μας ακολουθούν στο Google+

Τα social media(κοινωνικά μέσα ενημέρωσης) γίνονται σταδιακά τα νέα backlinks που θα επηρεάζουν ολοένα και περισσότερο την κατάταξη των ιστοσελίδων. Οι λόγοι που οδηγούν σε αυτό είναι πολλοί. Ο κόσμος μας

ασχολείται ολοένα και περισσότερο με τα κοινωνικά μέσα ενημέρωσης. Οι νέες γενιές έχουν μεγαλώσει με αυτά και τα εμπιστεύονται. Είναι πιο πιθανό να προτιμήσουμε μια ιστοσελίδα που την έχουν προτείνει οι ψηφιακοί μας φίλοι παρά οι ΜΑ. Συμπερασματικά η ενοποίηση των κοινωνικών μέσων ενημέρωσης και του SEO είναι μονόδρομος για αυτούς που θέλουν να έχουν υψηλή κατάταξη. Επιπρόσθετα η επίδραση των κοινωνικών μέσων ενημέρωσης ενισχύεται μέσα από αυτή την ενοποίηση γιατί θα τροφοδοτείται από την βελτίωση του SERP (Search Engine Results Placement) το οποίο θα δημιουργεί αυξημένη επισκεψιμότητα όχι μονάχα στην ιστοσελίδα αλλά και στις σελίδες κοινωνικής δικτύωσης του[38].

3.5 Λέξεις κλειδιά και που στοχεύουν (Keywords, target group)

Οι λέξεις κλειδιά αποτελούν το τρόπο με τον οποίο οι χρήστες θα μας βρουν μέσα από τις ΜΑ. Για το λόγο αυτό απαιτείται ιδιαίτερη προσοχή για την επιλογή των κατάλληλων λέξεων κλειδιών τα οποία θα είναι ανταγωνιστικά αλλά ταυτόχρονα θα πρέπει να ανταποκρίνονται και στο περιεχόμενο των σελίδων των οποίων αναφέρονται. Μεγάλη εξίσου προσοχή θα πρέπει να δείξουμε κατά την επιλογή ανταγωνιστικών λέξεων. Για παράδειγμα όταν η ίδια λέξη κλειδί χρησιμοποιείται από πολλές ιστοσελίδες οι οποίες είναι απόλυτα στοχευμένες σε αυτό τότε έχουν ποιοτική παρουσία στο χώρο του διαδικτύου με υψηλό PR(pagerank). Συνεπώς καλό είναι να προσπαθήσουμε να βρούμε κάποιο άλλο το οποίο θα είναι αρκετά ανταγωνιστικό αλλά ταυτόχρονα δεν έχει στοχευθεί ικανοποιητικά από πολλές ποιοτικές ιστοσελίδες. Αν για παράδειγμα οι πιο δημοφιλείς ιστοσελίδες που εμφανίζονται για τον προγραμματισμό arduino χρησιμοποιούν τη λέξη “arduino programming”, για τη δική μας σελίδα θα πρέπει να βρούμε μια λέξη ή φράση που να αποδίδει εξίσου το ίδιο. Αυτή θα μπορούσε να είναι “how to program an arduino” ή ακόμα και “arduino tutorials”. Φυσικά, το να ανταγωνιστεί κάποιος αυτές τις ιστοσελίδες που μπορεί εδώ και χρόνια να εμφανίζονται στην πρώτη σελίδα του Google για συγκεκριμένες λέξεις κλειδιά, σίγουρα θα πρέπει να περάσει αρκετός χρόνος (τουλάχιστον 6 μήνες) και θα πρέπει να κάνει πάρα πολλές ενέργειες SEO. Τα εργαλεία τα οποία μας βοηθούν στη εύρεση των σωστών λέξεων θα τα αναλύσουμε στο κεφάλαιο 4. Συνοψίζοντας

η επιλογή των σωστών λέξεων κλειδιών θα πρέπει να παρουσιάζει τα εξής χαρακτηριστικά:

- Οι λέξεις κλειδιά που επιλέγονται θα πρέπει να εμφανίζονται όσο το δυνατόν περισσότερο στις μηνιαίες αναζητήσεις των χρηστών και χαμηλό ανταγωνισμό.
- Η επιλογή των λέξεων πρέπει να είναι συνυφασμένη με το πεδίο το οποίο αναφέρεται στη σελίδα μας.
- Οι λέξεις με υψηλό ανταγωνισμό θα πρέπει να αποφεύγονται και στη θέση τους να χρησιμοποιούνται φράσεις κλειδιά.

3.6 Μειονεκτήματα SEO

Το SEO αποτελεί τη νούμερο πηγή αύξησης της επισκεψιμότητας για πολλές ιστοσελίδες, για αυτό είναι καλό να γνωρίζουμε και τα μειονεκτήματα του. Στη παρούσα ενότητα θα προσπαθήσουμε να τα απαριθμήσουμε ώστε να έχουμε μια ολοκληρωμένη άποψη.

1. Τα αποτελέσματα της δουλειάς μας πάνω στο SEO θα χρειαστούν κάποιο χρόνο μέχρι να καρποφορήσουν. Πιο συγκεκριμένα, θα πρέπει να λάβουμε υπόψη, το χρόνο τον οποίο θα χρειαστεί για να γίνουν οι κατάλληλες αλλαγές στην σελίδα καθώς επίσης και το χρόνο που θα χρειαστούν οι ΜΑ μέχρι να πάρουν το νέο περιεχόμενο και να το ευρετηριάσουν (indexing). Αν θέλουμε γρήγορα αποτελέσματα τότε ίσως ο δρόμος του SEO να μην είναι η καλύτερη λύση. Στο κεφάλαιο 6 θα δούμε τι άλλες επιλογές έχουμε και πως πρέπει να κινηθούμε.

2. Ως λογικό ακόλουθο μπορούμε να πούμε πως θα χρειαστεί επίσης ένα εύλογο χρονικό διάστημα μέχρι να γίνει η απόσβεση για την επένδυση που έχει γίνει πάνω στο SEO. Αυτό δεν είναι καθόλου καλό για τις μικρές επιχειρήσεις που χρειάζονται έσοδα σε σύντομο χρονικό διάστημα. Ο αγγλικός όρος που αντιπροσωπεύει τα παραπάνω είναι ROI, συντομογραφία του Return of Investments.

3. Το SEO απαιτεί κάποια αξιόλογη επένδυση ώστε να διασφαλιστεί μία θέση ανταγωνιστικότητας. Το SEO δεν αποτελεί κάποιου είδους δωρεάν εκστρατεία εκτός και αν μπορείς να παράξεις μόνος σου πρωτογενές περιεχόμενο καθώς και να προσελκύσεις ένα εύλογο αριθμό backlink. Συνήθως για να γίνουν αυτά χρειάζεται να προσληφθεί κάποιος αριθμός ατόμων ή εξωτερικών συνεργατών ώστε να καλυφθούν οι δραστηριότητες που απαιτούνται για τη παραγωγή σημαντικών αποτελεσμάτων.

4. Η εμφάνιση στη πρώτη σελίδα των ΜΑ πρέπει να αποτελεί το κύριο στόχο σου διότι σπάνια ένας μέσος χρήστης θα περιηγηθεί σε περισσότερες από μία σελίδες από τα αποτελέσματα που θα εμφανιστούν.

5. Καθώς κανείς δε γνωρίζει ακριβώς το τρόπο λειτουργίας των αλγορίθμων που τρέχουν πίσω από τις ΜΑ, δε μπορείς να ελέγξεις πόση κίνηση θα υπάρχει από όλη τη προσπάθεια βελτιστοποίησης που έχει καταβληθεί προηγουμένως.

6. Καθώς πολλά πράγματα σχετικά με το SEO μπορεί να αλλάξουν στο μέλλον όπως έγινε με τον ερχομό των panda και penguin, θα πρέπει να υπάρχουν συνεχείς ενημερώσεις. Η ιστοσελίδα θα πρέπει συνεχώς να είναι όσο το δυνατό φιλική στη ΜΑ χωρίς να δημιουργούνται προβλήματα στην εμπειρία του χρήστη (User Experience-UX) διότι μπορεί να υπάρχουν ποινές από αυτή επηρεάζοντας τη κατάταξη των σελίδων[33].

4 SEO tools- software

Σε αυτό το κεφάλαιο θα δούμε και θα εξηγήσουμε τα σημαντικότερα εργαλεία που υπάρχουν ώστε να μας βοηθήσουν στην βελτιστοποίηση της ιστοσελίδας μας. Τα εργαλεία αυτά θα μπορούσαμε να τα κατατάξουμε σε κατηγορίες ως εξής:

- Εργαλεία αναζήτησης λέξεων κλειδιών.
- Εργαλεία ελέγχου της ιστοσελίδας μας.
- Εργαλεία για την βελτιστοποίηση της ιστοσελίδας μας.
- και άλλα εργαλεία που μπορούν να χρησιμοποιηθούν για να μας βοηθήσουν.

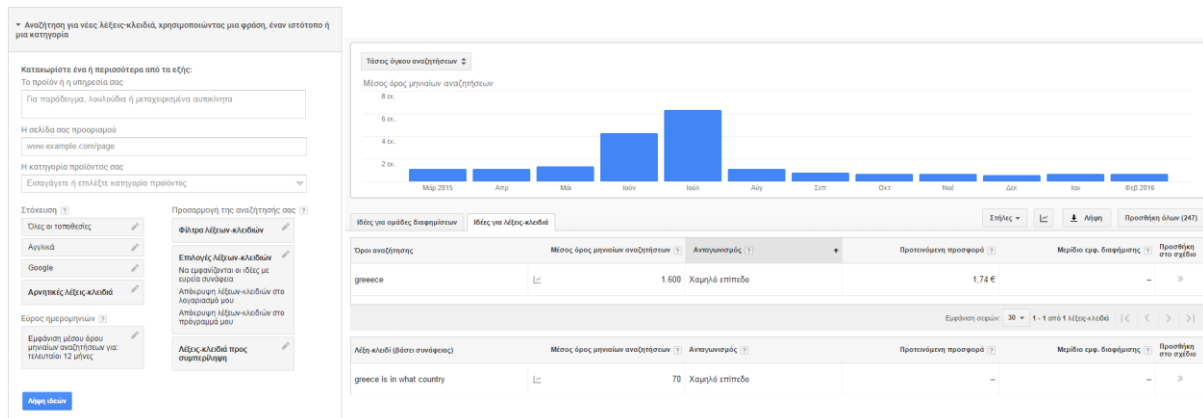
4.1 Εργαλεία αναζήτησης λέξεων κλειδιών

Η αναζήτηση λέξεων είναι μία βασική μέθοδος που χρησιμοποιούν επαγγελματίες για να βρουν πως οι χρήστες πραγματοποιούν τις αναζητήσεις τους στις ΜΑ. Η εύρεση των σωστών λέξεων κλειδιών αποτελεί μια συνεχή εργασία καθώς μπορεί με το χρόνο να μεταβληθεί ο τρόπος με τον οποίο ψάχνουν οι χρήστες στο διαδίκτυο. Η σωστή επιλογή αυτών αποτελεί καθοριστικό παράγοντα στη κατάταξη της σελίδας. Τα σημαντικότερα εργαλεία είναι :

Google adwords: keyword planner

Για τη χρησιμοποίηση του εργαλείου adwords που προσφέρει η google το μόνο που χρειάζεται είναι η δημιουργία ενός λογαριασμού με έναν λογαριασμό gmail. Οι υπηρεσίες που προσφέρει είναι δωρεάν και χρησιμοποιείται κυρίως για την έρευνα λέξεων κλειδιών αλλά και λήψη στατιστικών στοιχείων και προβλέψεων επισκεψιμότητας ενώ μπορεί να χρησιμοποιηθεί και για τη δημιουργία διαφημίσεων. Επίσης υπάρχουν αρκετά φίλτρα τα οποία κάνουν την αναζήτηση μας πιο στοχευμένη και προτείνονται και άλλες λέξεις κλειδιά ανάλογα με το τι έχουμε πληκτρολογήσει.

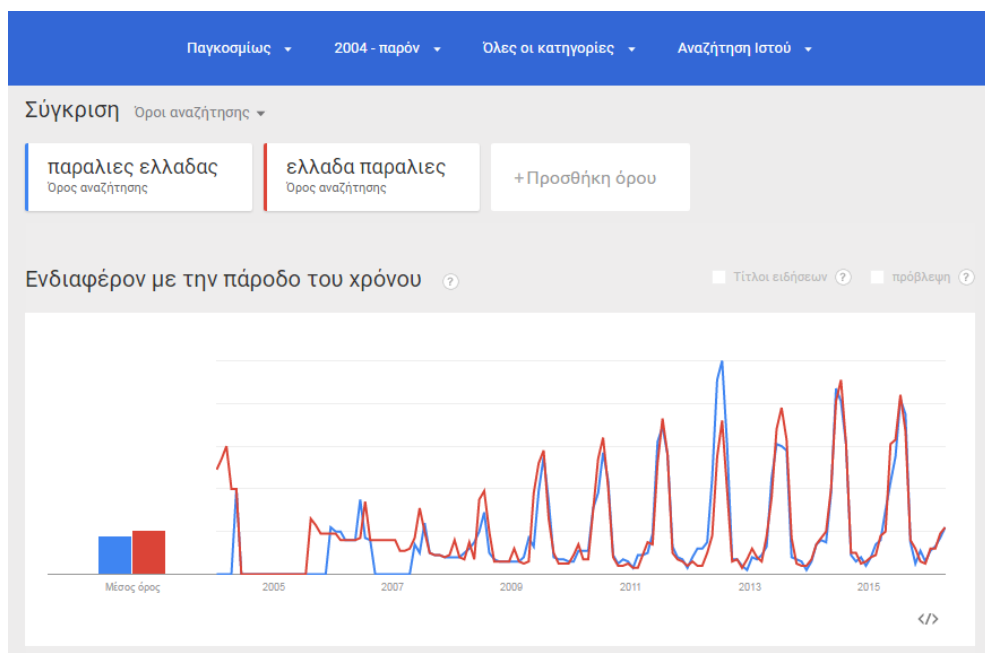
Τεχνικές, μέθοδοι και κανόνες για τη βελτιστοποίηση ανάπτυξης ιστοσελίδων



Εικόνα 4.1: Keyword planner της google.

Google trends

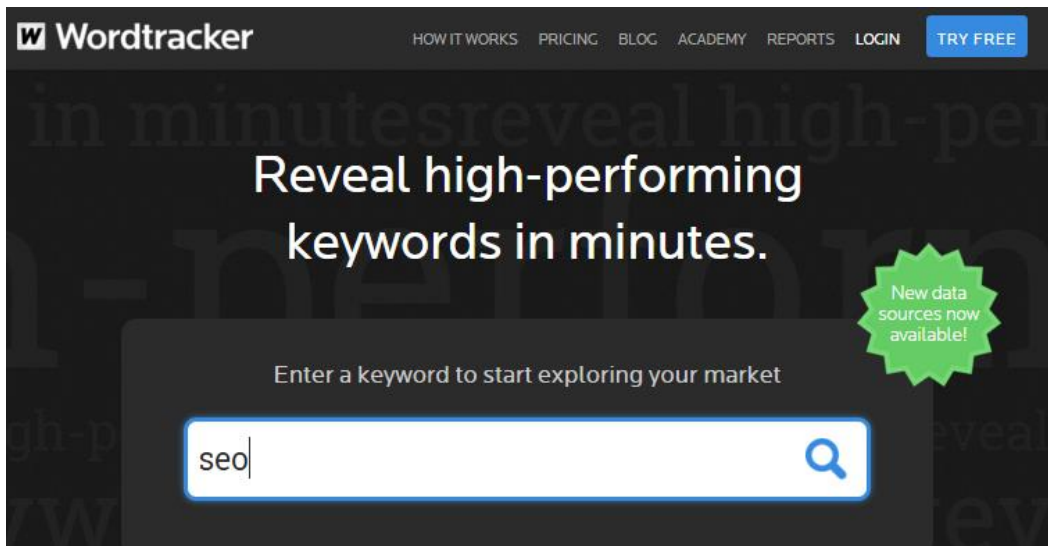
Το google trends αποτελεί μια ακόμα δωρεάν υπηρεσία της google η οποία μας δίνει τη δυνατότητα να δούμε τη δημοτικότητα των λέξεων κλειδιών. Επίσης μας επιτρέπει να συγκρίνουμε διαφορετικές λέξεις κλειδιά καθώς και τη χρονική περίοδο που είναι πιο δημοφιλή. Τέλος παρέχονται ειδικά φίλτρα ώστε να γίνουν πιο στοχευμένες αναζητήσεις παρέχοντας πιο ακριβή αποτελέσματα.



Εικόνα 4.2: Google trends.

Wordtracker

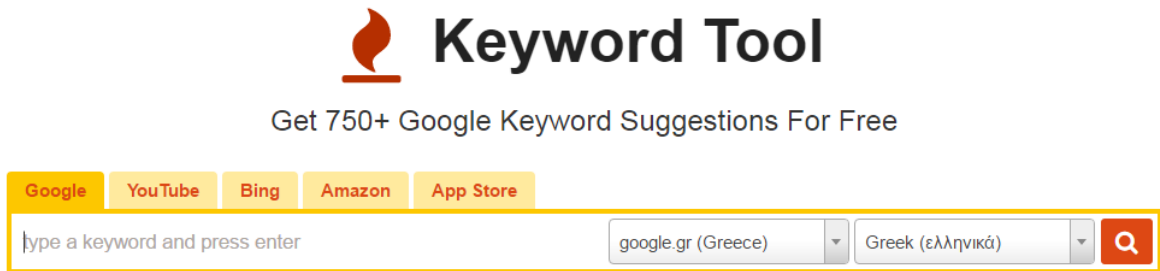
Το συγκεκριμένο εργαλείο παρέχει μία τεράστια βάση δεδομένων καθώς για μία λέξη μπορεί να προσφέρει μέχρι και μερικές εκατοντάδες σχετικές λέξεις. Επιπλέον προσφέρει τη δυνατότητα αναζήτησης για διάφορες λέξεις κλειδιά ταυτόχρονα, ενώ διαθέτει και φίλτρα τα οποία εξειδικεύουν περισσότερο την αναζήτηση μας.



Εικόνα 4.3: Wordtracker.

Keyword tool

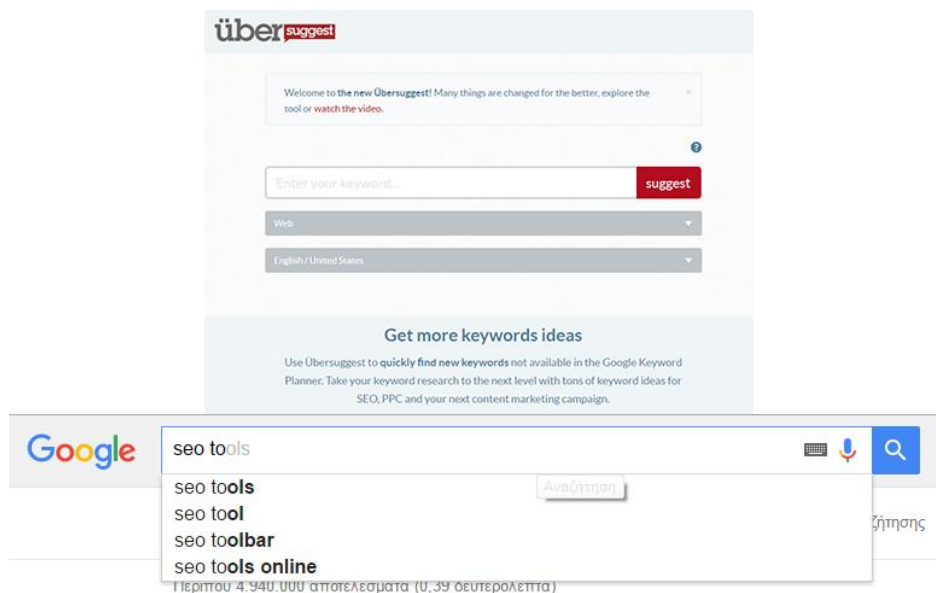
Ένα πολύ δυνατό εργαλείο που μπορούμε να βρούμε στη τοποθεσία keywordtool.io το οποίο είναι όμως διαθέσιμο επί πληρωμή. Μπορεί να χρησιμοποιηθεί αντί του google adwords και προσφέρει μία πληθώρα από χαρακτηριστικά. Προσφέρει τη δυνατότητα εύρεσης λέξεων κλειδιών και σε άλλες τοποθεσίες και MA όπως youtube και bing και ταυτόχρονα προτείνει άλλα τα οποία θα μπορούσαν να χρησιμοποιηθούν.



Εικόνα 4.4: Keyword Tool.

Google suggest- Ubersuggest

Τα συγκεκριμένα εργαλεία είναι παρόμοια. Όταν ο χρήστης πληκτρολογεί κάτι στο πεδίο αναζήτησης(λέξη ή φράση) εμφανίζονται σε αυτό προτάσεις οι οποίες βασίζονται στους πιο πολυαναζητημένους όρους.

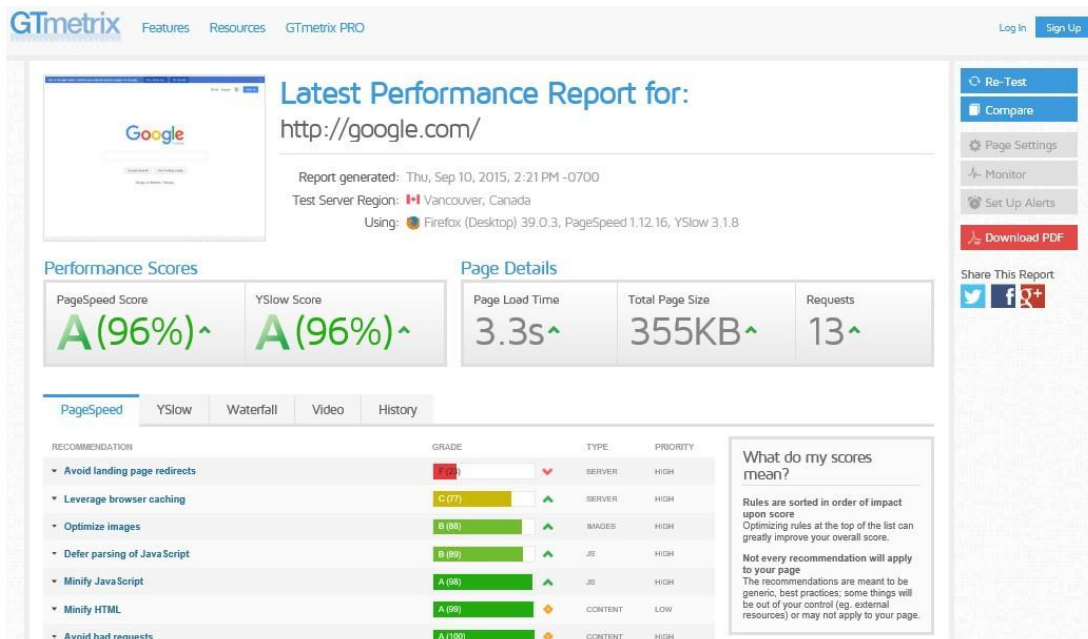


Εικόνα 4.5: Ubersuggest και google suggest αντίστοιχα.

4.2 Εργαλεία ελέγχου ιστοσελίδας

GtMetrix

Πολύ καλό εργαλείο για την μέτρηση της ταχύτητας της ιστοσελίδας σας. Μας δίνει το χρόνο που χρειάζεται για να φορτώσει η τελευταία καθώς και τις μετρήσεις από δύο αξιόπιστα εργαλεία, το Page Speed και το YSlow. Από τα δύο τελευταία προκύπτουν και προτάσεις για την βελτίωση της ταχύτητας της ιστοσελίδας μας.



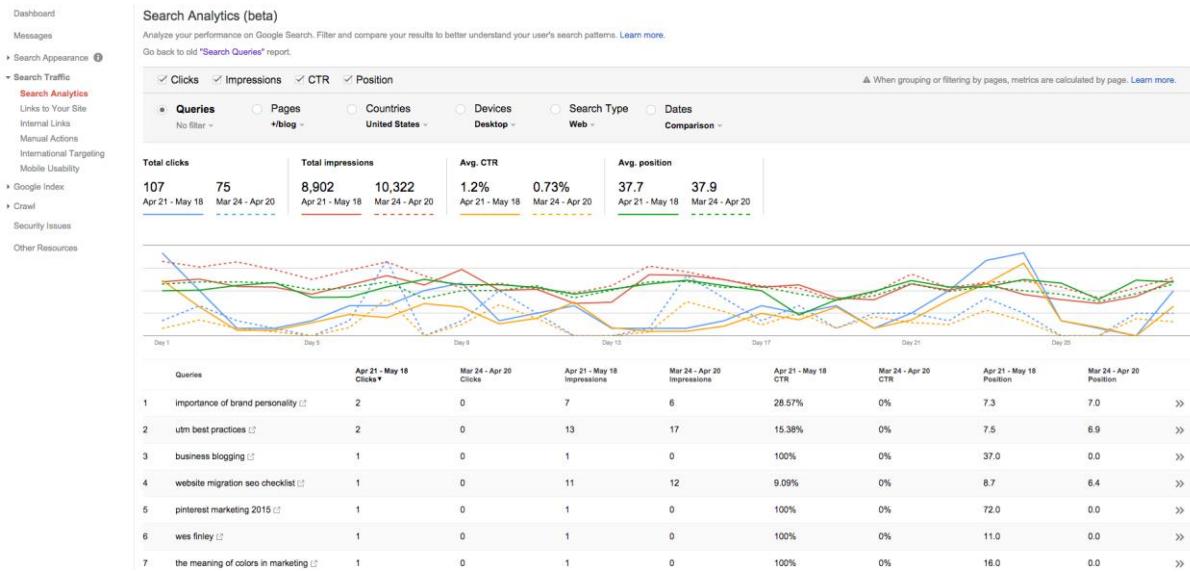
Εικόνα 4.6: Gtmetrix.

Google Search Console(Google webmaster tools)

Πολλοί μπορεί να το γνωρίζουν με την ονομασία webmaster tools, όμως η google άλλαξε το όνομα της σε search consoles στα μέσα του 2015. Αποτελεί μια δωρεάν υπηρεσία στην οποία δίνεται η δυνατότητα υποβολής και ελέγχου των sitemaps και του αρχείου robots.txt. Παράλληλα προσφέρει μια λίστα με τις εσωτερικές και εξωτερικές συνδέσεις καθώς και μια λίστα με τις συνδέσεις spam. Γνωστοποιεί τους δείκτες CTR(click threw rate) καθώς και πιθανές τιμωρίες σε

Τεχνικές, μέθοδοι και κανόνες για τη βελτιστοποίηση ανάπτυξης ιστοσελίδων

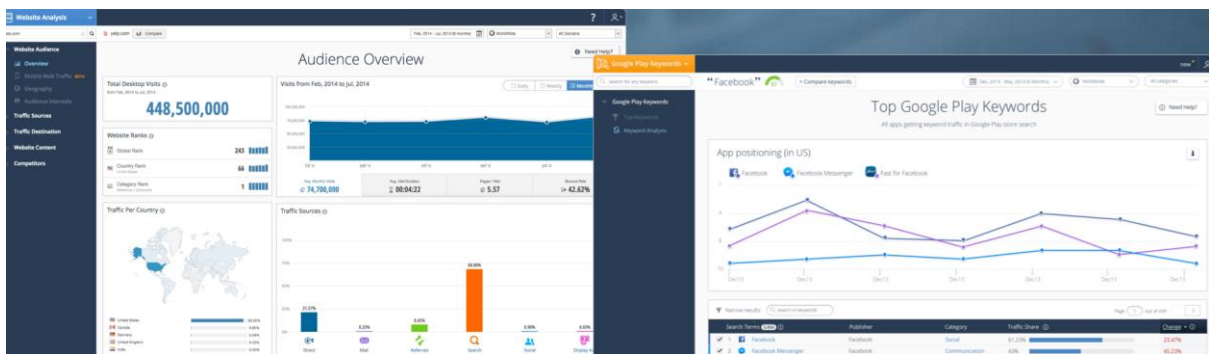
περίπτωση που έχουν χρησιμοποιηθεί αθέμιτοι τρόποι SEO. Συνοψίζοντας μας πληροφορεί με αναλυτικά δεδομένα για ότι συμβαίνει στην ιστοσελίδα μας[39].



Εικόνα 4.7: Google search console.

Similar web

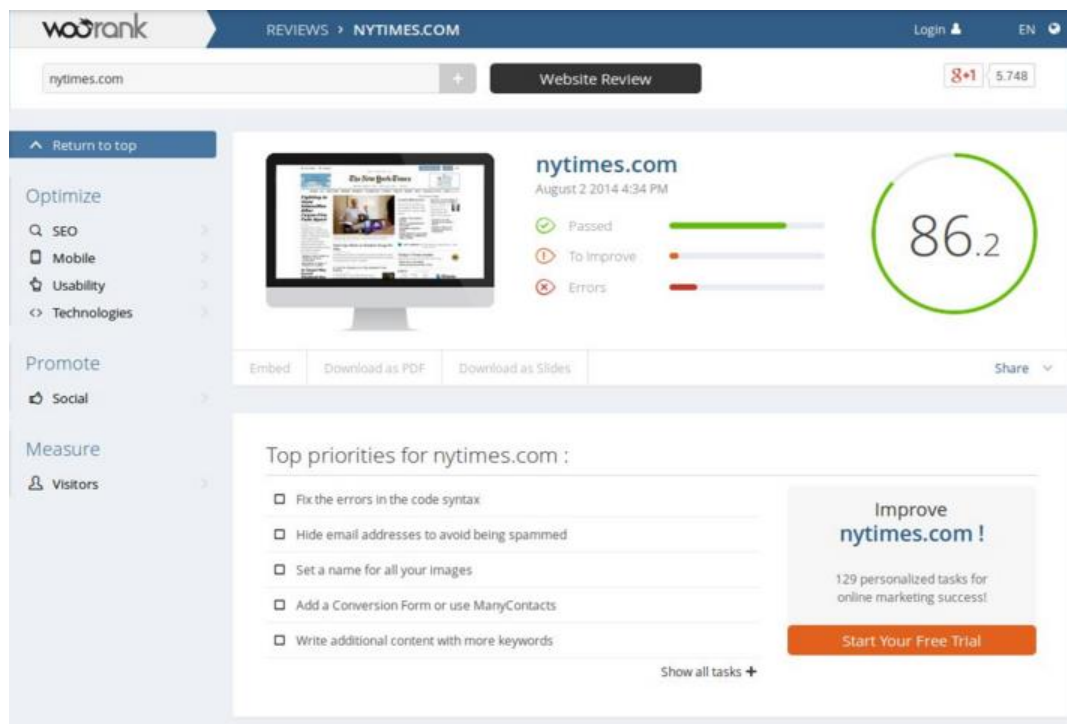
Το εργαλείο αυτό δείχνει τη κατάταξη μιας ιστοσελίδας σε παγκόσμιο και εθνικό επίπεδο. Μοιάζει με το alexa, όμως το similarweb παρέχει περισσότερες πληροφορίες. Αυτές είναι σχετικά με το αν διαφημίζεται μια ιστοσελίδα, τη κατάταξη για τη κατηγορία της, την ύπαρξη εφαρμογής για κινητά τηλέφωνα και τη δυνατότητα σύγκρισης με ανταγωνιστές (ιστοσελίδες).



Εικόνα 4.8: Similarweb.

Woorank

Χρήσιμο εργαλείο για την ανάλυση μιας ιστοσελίδας. Το δυνατό του σημείο είναι ότι κάνει αρκετούς ελέγχους ταυτόχρονα που αφορούν και onsite και offsite optimization. Το αδύνατο του σημείο είναι ότι τα αποτελέσματα ορισμένων ελέγχων δεν είναι ιδιαίτερα ακριβή. Παρόλα αυτά είναι επαρκέστατο για μια πρώτη αξιολόγηση μιας ιστοσελίδας[40].

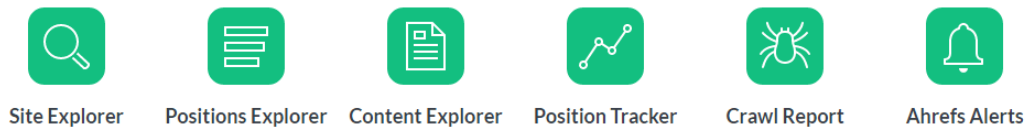


Εικόνα 4.9: Woorank.

Ahrefs

Αποτελεί ένα από τα καλύτερα εργαλεία SEO για τον έλεγχο της ποιότητας των backlink της ιστοσελίδας μας αλλά και των ανταγωνιστών μας. Επίσης δείχνει τη κίνηση ιστοσελίδων ανταγωνιστών, πληροφορεί για τα backlink της ιστοσελίδας μας τα οποία είναι πιο δυνατά και πως μπορούμε να χτίσουμε τη στρατηγική μας γύρω από αυτά. Παρέχεται επί πληρωμή αλλά και δωρεάν.

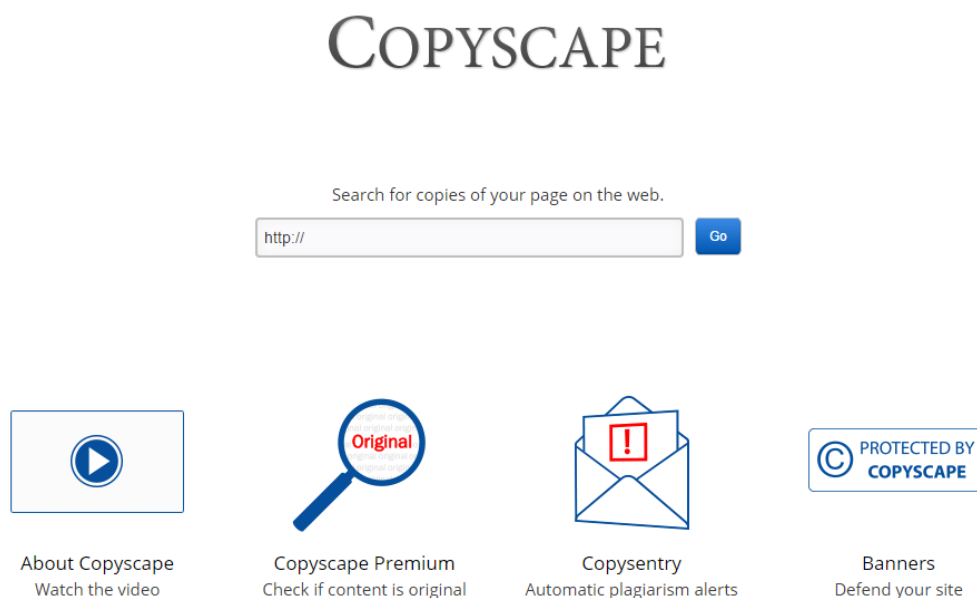
The tools you get with Ahrefs



Εικόνα 4.10: Ahrefs.

Copyscape

Το copyscape χρησιμοποιείται για να ελέγξει αν έχει αντιγραφεί περιεχόμενο της ιστοσελίδας μας. Για να πιστοποιηθεί ότι το περιεχόμενο μας είναι το αυθεντικό χρειάζεται να συνδεθούμε με το google+ και να χρησιμοποιήσουμε την υπηρεσία authorship. Επειδή στο διαδίκτυο η κλοπή περιεχομένου είναι πολύ συνηθισμένη η google επιβραβεύει τους δημιουργούς πρωτότυπων περιεχομένων δίνοντας τους καλύτερη κατάταξη.

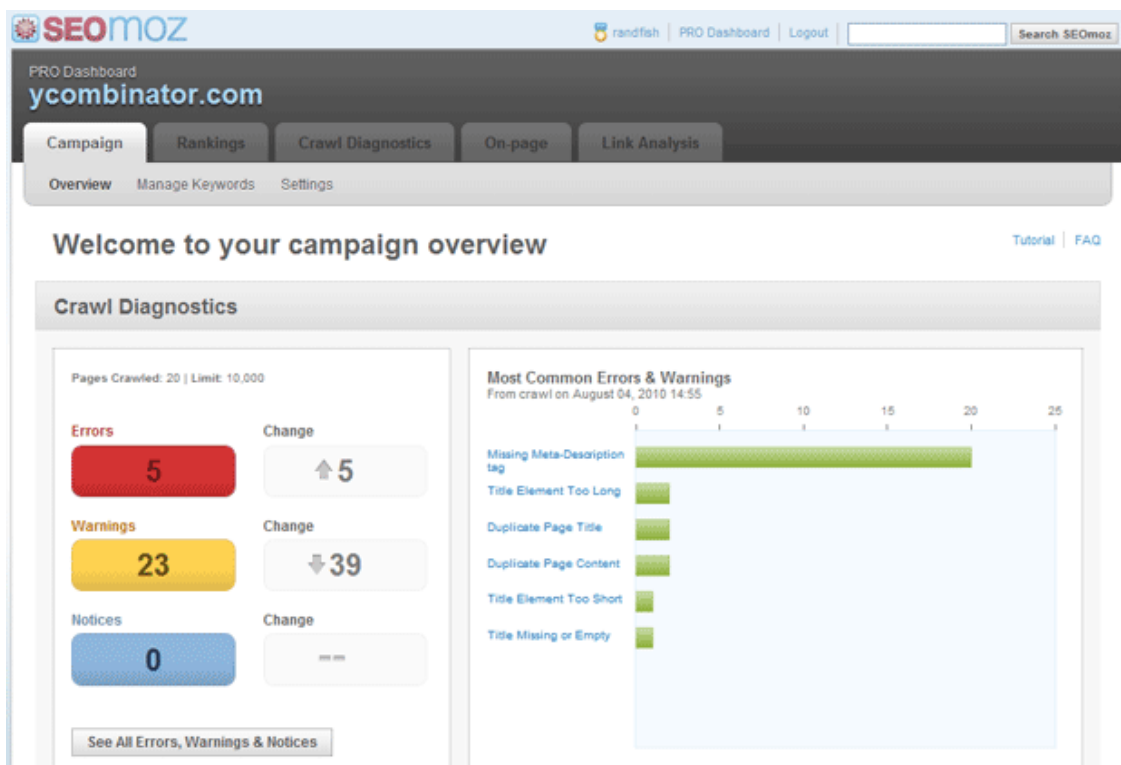


Εικόνα 4.11: Copyscape.

4.3 Εργαλεία βελτιστοποίησης

Μoz ή SEOmoz για τους παλιότερους

Το moz περιέχει μια σειρά από εργαλεία τα οποία μας βοηθούν στην κατάταξη της ιστοσελίδας μας στις ΜΑ. Έχει αρκετά φιλικό περιβάλλον και προτείνεται για νέους χρήστες. Τέλος παρέχει συμβουλές για βελτιστοποίηση στο on-site seo κομμάτι. Έχει εργαλεία τα οποία ερευνούν για τις πιο δημοφιλείς αναζητήσεις, συγκρίνει την αξιοπιστία ιστοσελίδων μέσω του εργαλείου moztrust, έχει δικό της εργαλείο για τον υπολογισμό του pagerank γνωστό ως mozrank και άλλα.



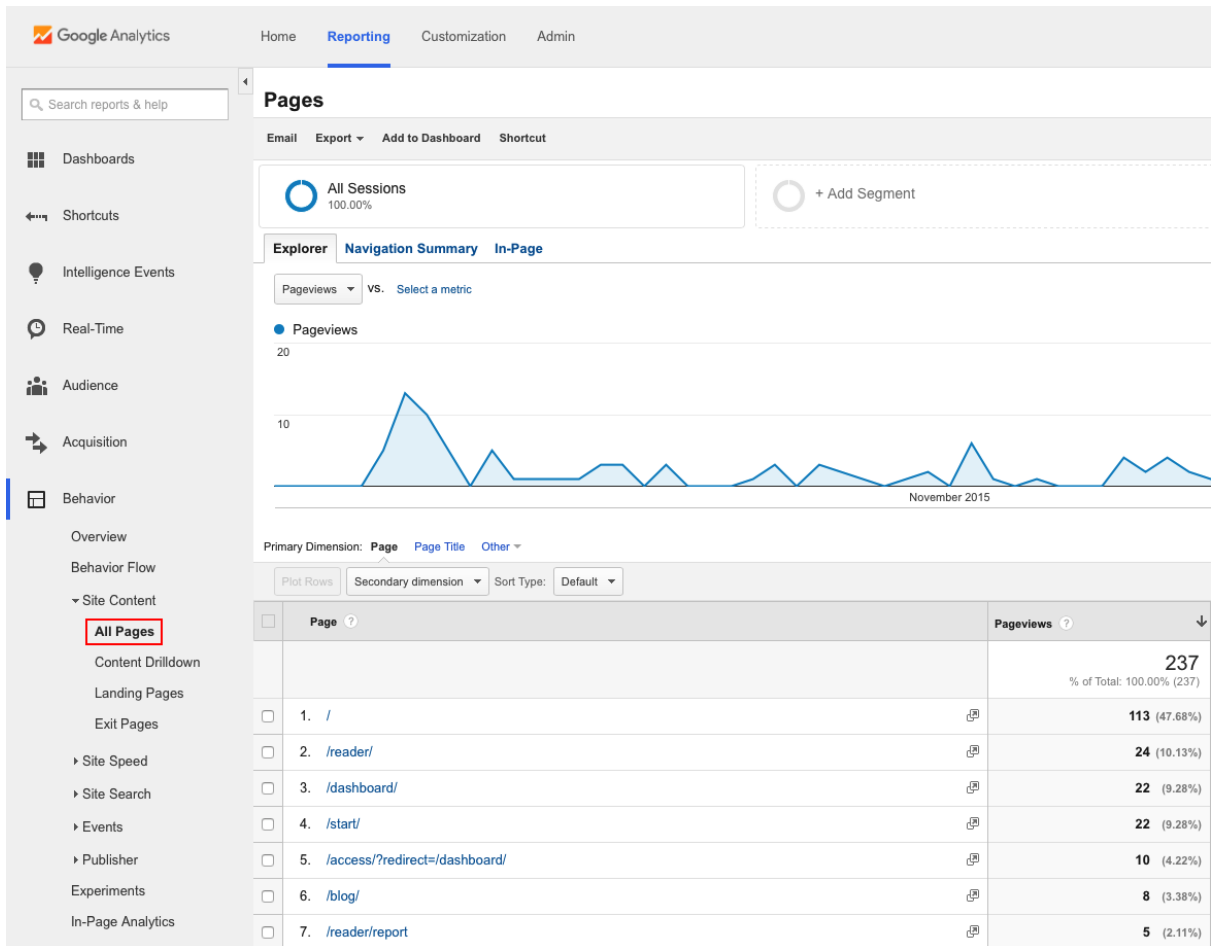
Εικόνα 4.12: Moz[41].

Google analytics

Το συγκεκριμένο εργαλείο παρέχεται δωρεάν από τη google και δίνει τη δυνατότητα στο χρήστη να εντοπίσει το περιεχόμενο της σελίδας του το οποίο

Τεχνικές, μέθοδοι και κανόνες για τη βελτιστοποίηση ανάπτυξης ιστοσελίδων

θεωρείται πιο δημοφιλής. Επίσης παρέχει αναλυτικές αναφορές με την επισκεψιμότητα των χρηστών στην ιστοσελίδα μας.



Εικόνα 4.13: Google analytics.

Google pagespeed

Το google pagespeed είναι μια οικογένεια από εργαλεία της google τα οποία έχουν σχεδιαστεί για να βελτιώσουν τις επιδόσεις της ιστοσελίδας. Πρώτη φορά παρουσιάστηκε το 2010 και αποτελείται από 4 βασικά στοιχεία(pagespeed module, page speed insights, page speed service και pagespeed chrome). Όλα αυτά εξυπηρετούν στον εντοπισμό των λαθών μιας ιστοσελίδας για την καλύτερη και ταχύτερη επίδοση της.

4.4 Άλλα εργαλεία

Broken link checker

Πρακτικό και εύκολο στη χρήση του εργαλείο το οποίο εντοπίζει τους συνδέσμους που δε λειτουργούν (broken links) της ιστοσελίδας μας. Εντοπίζει το URL στο οποίο βρίσκεται το broken link καθώς και την θέση του στον κώδικα της ιστοσελίδας[40].

Cloaking checker

Ανήκει στην κατηγορία των εργαλείων που ελέγχει αν η ιστοσελίδα που μας ενδιαφέρει χρησιμοποιεί μια παράνομη τεχνική βελτιστοποίησης ιστοσελίδων με την ονομασία cloaking[40].

Google copyright removal

Εργαλείο που βρίσκεται στα Google Webmaster Tools και το οποίο μας επιτρέπει να αναφέρουμε στην Google τις ιστοσελίδες που μας έχουν αντιγράψει το περιεχόμενό μας[40].

Google pagerank center

Ελέγχει το PageRank (PR), δηλαδή τη βαθμολογία που θεωρητικά δίνει η Google στην ιστοσελίδα μας. Από το εργαλείο μπορούμε επίσης να δούμε και το PR των ανταγωνιστών μας ή των ιστοσελίδων που σκοπεύουμε να δημιουργήσουμε backlinks. Βέβαια το πραγματικό PR που δίνει η Google δεν το γνωρίζει κανείς, το εργαλείο αυτό δίνει μια εκτίμηση για αυτό[40].

Robots.txt checker

Ελέγχει το robots.txt αρχείο της ιστοσελίδας μας για να εντοπίσει πιθανά λάθη τα οποία εμποδίζουν την πρόσβαση των MA σε αυτή[40].

Whatsmyserp

Εξαιρετικό εργαλείο για να δούμε την κατάταξη της ιστοσελίδας μας για τις λέξεις-κλειδιά που μας ενδιαφέρουν. Ο έλεγχος της κατάταξης πραγματοποιείται διαδικτυακά οπότε δεν έχουμε τον κίνδυνο να μας μπλοκάρει προσωρινά η Google τις αναζητήσεις. Επίσης μπορούμε ταυτόχρονα να δούμε και την κατάταξη των βασικών μας ανταγωνιστών[40].

Topsy

Εντοπίζει δημοφιλές περιεχόμενο στο Twitter για τις λέξεις-κλειδιά που μας ενδιαφέρουν και μας βοηθάει να δημιουργήσουμε το δικό μας. Βάζουμε την λέξη-κλειδί που μας ενδιαφέρει και επιλέγουμε συνδέσμους. Με τον τρόπο αυτό θα μας εμφανιστούν τα περιεχόμενα που έχουν τα περισσότερα backlinks και συνεπώς βασιζόμενοι σε αυτά και αναπτύσσοντας το δικό μας πρωτότυπο περιεχόμενο θα έχουμε περισσότερες πιθανότητες να δημιουργήσουμε φυσικά backlinks προς την ιστοσελίδα μας[40].

Reduce bounce rate

Εξαιρετικό plugin το οποίο συμβάλλει στην σημαντική μείωση του ποσοστού εγκατάλειψης(bounce rate) της ιστοσελίδας μας. Τα Google Analytics μετράνε ως εγκατάλειψη(bounce) τους επισκέπτες της ιστοσελίδας οι οποίοι βλέπουν μονάχα μια σελίδα της και μετά φεύγουν, ανεξάρτητα από το χρόνο που έχουν μείνει σε αυτή ή τις ενέργειες που έχουν γίνει σε αυτή. Με το συγκεκριμένο plugin αν ο επισκέπτης της ιστοσελίδας μας μείνει σε μια σελίδα της περισσότερο από κάποιο χρονικό διάστημα (προεπιλεγμένο είναι τα 10 sec) ή κάνει κάποια ενέργεια όπως είναι το scroll down τότε δεν θα καταγραφεί ως bounce ακόμα και αν δεν επισκεφθεί άλλη σελίδα. Με τον τρόπο αυτό βελτιώνονται τα στατιστικά στοιχεία της ιστοσελίδας μας και συνεπώς έχουμε υψηλότερη βαθμολογία από την Google[40].

5 SEO για χρήστες κινητών τηλεφώνων

Καθώς τα smartphone έχουν κατακλύσει την αγορά, όλο και περισσότεροι χρήστες τα χρησιμοποιούν για να συνδεθούν στο διαδίκτυο. Οι κυριότεροι λόγοι που οι χρήστες τους τα χρησιμοποιούν είτε για να βρουν πληροφορίες είτε για διασκεδάσουν τον ελεύθερο τους χρόνο. Κύριο χαρακτηριστικό για τους χρήστες που περιηγούνται μέσω των κινητών είναι ότι απαιτούν οι αναζητήσεις στους να είναι όσο το δυνατόν πιο γρήγορες. Στο συγκεκριμένο κεφάλαιο θα εξετάσουμε το seo για τις συσκευές των κινητών τηλεφώνων σύμφωνα με τα πρότυπα της google.

5.1 Κατασκευή mobile friendly ιστοσελίδων σύμφωνα με τη google

Οι 3 βασικές τεχνικές οι οποίες χρειάζονται να εφαρμοστούν σε μία ιστοσελίδα ώστε να γίνεται αντιληπτή και να βοηθήσει τις MA να τη κατανοήσουν είναι :

- **Responsive web design:** Συνίσταται ως πρότυπο σχεδιασμού από τη google. Ανεξάρτητα από τη συσκευή του χρήστη(desktop, mobile κ.α) παρέχεται ο ίδιος κώδικας HTML. Το τελικό αποτέλεσμα το οποίο εμφανίζεται στη συσκευή εξαρτάται από το μέγεθος της οθόνης της.
- **Dynamic Serving:** Ανεξάρτητα από τη συσκευή που θα φορτώνει η σελίδα θα πρέπει η σελίδα να είναι σε θέση να δημιουργεί για το ίδιο URL διαφορετικές εκδόσεις HTML. Οι εκδόσεις HTML βασίζονται στο τι γνωρίζει ο server για το φυλλομετρητή (browser) του χρήστη.
- **Ξεχωριστά URL:** Η συγκεκριμένη ρύθμιση προσπαθεί να ανιχνεύσει τη συσκευή του χρήστη ώστε να τον ανακατευθύνει στη συνέχεια στη κατάλληλη σελίδα.

Είναι επίσης σημαντικό να κατανοήσουμε τις διάφορες συσκευές στις οποίες απευθυνόμαστε:

- **Mobile:** Με τη χρήση του όρου mobile απευθυνόμαστε στις έξυπνες συσκευές κινητών τηλεφώνων (smartphone) Android, Iphone και Windows Phone. Οι browsers των κινητών μοιάζουν με αυτές των σταθερών υπολογιστών (desktops) παρόλο που οι οθόνες τους είναι μικρότερες. Οι ίδιες μπορούν να κάνουν render ένα ευρύ σύνολο της HTML 5.
- **Tablet:** Οι συγκεκριμένες συσκευές αποτελούν ξεχωριστή κατηγορία. Γενικά όταν μιλάμε για mobile συσκευές δε συμπεριλαμβάνουμε και τις συσκευές tablet. Αυτές έχουν μεγαλύτερες οθόνες και οι χρήστες αναμένουν να φαίνονται οι σελίδες όπως στον επιτραπέζιο υπολογιστή.
- **Multimedia:** Οι συγκεκριμένες τηλεφωνικές συσκευές μπορούν να παράξουν σελίδες οι οποίες είναι σύμφωνα με τις προδιαγραφές XHTML, υποστηρίζουν HTML 6, Javascript/ECMAScript αλλά μπορεί να μην υποστηρίζουν κάποιες επεκτάσεις(extensions) API της HTML5. Η συγκεκριμένη περιγραφή απευθύνεται σε browsers 3G συσκευές οι οποίες δεν είναι όμως smartphone.
- **Feature:** Οι τηλεφωνικές συσκευές feature δεν έχουν την ικανότητα να προβάλλουν σελίδες σύμφωνα με το πρότυπο της HTML παραμόνο με cHTML(iMode), WML XHTML-Mp.

5.1.1 Responsive web design

Ο σχεδιασμός διαδραστικών σελίδων στο διαδίκτυο(responsive web design) είναι η διαδικασία στην οποία ο server αποστέλλει τον ίδιο κώδικα HTML σε όλες τις συσκευές ενώ η CSS χρησιμοποιείται για να αλλάξει την εμφάνιση της σελίδας για κάθε συσκευή. Ο αλγόριθμος που χρησιμοποιείται πρέπει να είναι σε θέση να μπορεί να ανιχνεύσει αυτόματα τα στοιχεία της σελίδας (CSS, Javascript και εικόνες).



Εικόνα 5.1: Διανομή html σε διάφορες συσκευές[26].

Για να ειδοποιήσεις το φυλλομετρητή (browser) ώστε να διανέμει τη σωστή σελίδα στις διάφορες συσκευές θα πρέπει να χρησιμοποιήσεις το meta tag με τη παράμετρο name ως εξής:

```
<meta name="viewport" content="width=device-width, initial-scale=1.0">
```

Εικόνα 5.2: Meta tag[26].

Το meta tag δίνει στο φυλλομετρητή τις οδηγίες ώστε να προσαρμόσει τις διαστάσεις της σελίδας στο πλάτος της κάθε συσκευής. Όταν αυτό λείπει οι φυλλομετρητές χρησιμοποιούν τις προκαθορισμένες διαστάσεις προβολής της σελίδας όπως στον επιτραπέζιο υπολογιστή. Έπειτα οι φυλλομετρητές προσπαθούν να προβάλλουν το περιεχόμενο της σελίδας πιο καλαίσθητο αυξάνοντας το μέγεθος της γραμματοσειράς και αυξομειώνουν το περιεχόμενο ώστε να ταιριάζει στις διαστάσεις της οθόνης ή θα εμφανίζει μόνο μέρος του περιεχομένου εντός της οθόνης. Για τους χρήστες αυτό σημαίνει πως το μέγεθος της γραμματοσειράς που θα εμφανίζεται στην οθόνη μπορεί να είναι δυσανάγνωστο, με αποτέλεσμα να χρειαστεί να το μεγενθύνουν ώστε να διακρίνουν το περιεχόμενο της σελίδας.



Εικόνα 5.3: Non responsive vs responsive[26].

Οι λόγοι που εξυπηρετούν το responsive design είναι οι ακόλουθοι:

- Είναι ευκολότερο για τους χρήστες να μοιράζονται τις ιστοσελίδες χρησιμοποιώντας ένα url.
- Βοηθάει τη ΜΑ ώστε να εκχωρήσει με ακρίβεια τις ιδιότητες της σελίδας στο ευρετήριο (index).
- Για τη συντήρηση διάφορων σελίδων με το ίδιο περιεχόμενο απαιτείται λιγότερη συντήρηση.
- Μειώνει τη πιθανότητα λαθών στις ιστοσελίδες για τις διάφορες συσκευές.
- Δεν χρειάζεται ανακατεύθυνση για χρήστες που έχουν βελτιστοποιημένες τις σελίδας με αποτέλεσμα να μειώνεται ο χρόνος φόρτωσης της σελίδας.
- Μειώνονται οι πόροι που απαιτούνται από την αράχνη(bot) ώστε να σαρώσει (crawl) τη σελίδα. Οι διαδραστικές (responsive) σελίδες έχουν το πλεονέκτημα ότι δε σαρώνονται πολλές φορές για κάθε είδους συσκευή. Επίσης βοηθάει στο να διατηρείται το περιεχόμενο της σελίδας μόνιμως ενημερωμένο.

5.1.2 Δυναμική εξυπηρέτηση(Dynamic Serving)

Η δυναμική εξυπηρέτηση είναι μια εγκατάσταση όπου ο διακομιστής απαντά με διαφορετικό HTML και CSS για το ίδιο URL ανάλογα με τη σελίδα που ζητάει ο χρήστης. Δεδομένου ότι δεν είναι άμεσα εμφανές η αλλαγή του κώδικα της HTML της σελίδας για τους χρήστες των κινητών συσκευών ο διακομιστής(server) ζητάει από το bot να σαρώσει(crawl) τη σελίδα ώστε να ανιληφθεί το περιεχόμενο για τις κινητές συσκευές. Η υπόδειξη αυτή καταφαίνεται μέσα στον HTTP header.



Εικόνα 5.4: Δυναμική εξυπηρέτηση του ίδιου URL για διαφορετικές συσκευές[27].

5.1.2.1 Λειτουργία του vary http header

Ο vary http header έχει 2 βασικές και χρήσιμες λειτουργίες:

- Υπενθυμίζει στους διακομιστές που αποθηκεύουν προσωρινά τις ιστοσελίδες(caching servers)των διαφόρων ISP ότι πρέπει να σκεφτούν αν θα διαμοιράσουν τη σελίδα από cache ή όχι. Χωρίς αυτή τη δυνατότητα που προσφέρει ο http header υπάρχει η πιθανότητα μια προσωρινά αποθηκευμένη σελίδα να διαμοιραστεί σε λάθος συσκευή. Για παράδειγμα, μια σελίδα που προορίζεται για τις κινητές συσκευές θα μπορούσε να σερβιριστεί στους χρήστες επιτραπέζιων υπολογιστών και το αντίστροφο.
- Βοηθάει το ρομπότ(googlebot) να ανακαλύψει γρηγορότερα το περιεχόμενο το οποίο προορίζεται για την βέλτιστη εμφάνιση στις κινητές συσκευές.

```
GET /page-1 HTTP/1.1
Host: www.example.com
(...rest of HTTP request headers...)

HTTP/1.1 200 OK
Content-Type: text/html
Vary: User-Agent
Content-Length: 5710
(... rest of HTTP response headers...)
```

Εικόνα 5.5: Παράδειγμα vary http header[27].

5.1.3 Ξεχωριστά URL

Για κάθε ιστότοπο που επισκεπτόμαστε στο διαδίκτυο μέσω του υπολογιστή εισερχόμαστε σε μία διεύθυνση URL. Για να κάνουμε τις σελίδες μας ακόμα πιο φιλικές για τις κινητές συσκευές θα πρέπει κάθε URL που επισκεπτόμαστε να έχει ένα αντίστοιχο URL για τις κινητές συσκευές. Αν για παράδειγμα έχουμε τη σελίδα www.example.com η οποία εξυπηρετεί τους χρήστες των επιτραπέζιων υπολογιστών, για τους χρήστες των κινητών συσκευών θα πρέπει να παρασχεθεί με τη διεύθυνση m.example.com. Η google δεν προτείνει κάποια συγκεκριμένη μορφή του URL, αρκεί να έχει τη πρόσβαση που χρειάζεται το ρομπότ(googlebot) της.



Εικόνα 5.6: Διαφορετικές διευθύνσεις URL για την εξυπηρέτηση χρηστών από επιτραπέζιο υπολογιστή και κινητές συσκευές[28].

Η google επισημαίνει τους εξής σχολιασμούς για να βοηθηθούν οι αλγόριθμοι που χρησιμοποιούνται για να καταλάβουν τις διευθύνσεις URL των κινητών.

1. Στη σελίδα που προορίζεται για τους επιτραπέζιους υπολογιστές να προστεθεί η ετικέτα(tag), rel="alternate" το οποίο παραπέμπει στη σελίδα των κινητών συσκευών. Αυτό βοηθάει το ρομπότ(googlebot) να βρει ευκολότερα τη σελίδα για τη κινητή συσκευή.
2. Στη σελίδα που προορίζεται για τις κινητές συσκευές να προστεθεί η ετικέτα (tag) rel="canonical" η οποία θα δείχνει στη σελίδα για τους επιτραπέζιους υπολογιστές.

Αυτός ο σχολιασμός συμβαίνει για να υποστηρίξουν 2 μεθόδους: μέσα στην HTML της σελίδας καθώς και στο sitemap. Το sitemap αποτελεί ένα αρχείο στο οποίο είναι καταγεγραμμένες όλες οι διευθύνσεις(URL) της ιστοσελίδας και βοηθάει αρκετές MA στην οργάνωση του περιεχομένου αυτών. Επίσης το sitemap περιέχει σημαντικά μεταδεδομένα(metadata) σχετικά με τις διευθύνσεις(URL) που είναι καταχωρημένες και πληροφορίες που αφορούν τα βίντεο, τις εικόνες ακόμα και το περιεχόμενο για κινητές συσκευές. Έστω ότι έχουμε τις <http://example.com/page-1> και <http://m.example.com/page-1> σελίδες για τον επιτραπέζιο υπολογιστή και τις κινητές συσκευές αντίστοιχα. Οι σχολιασμοί για την HTML στους επιτραπέζιους υπολογιστές και στις κινητές συσκευές:

```
<link rel="alternate" media="only screen and (max-width: 640px)"  
href="http://m.example.com/page-1">
```

Εικόνα 5.7: Σχολιασμός(annotation) για την www.example.com/page-1 [28].

```
<link rel="canonical" href="http://www.example.com/page-1">
```

Εικόνα 5.8: Σχολιασμός(annotation) για την m.example.com/page-1 [28].

Ενώ το sitemap για τους υπολογιστές και τις κινητές συσκευές πρέπει να είναι:

```
<?xml version="1.0" encoding="UTF-8"?>
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"
  xmlns:xhtml="http://www.w3.org/1999/xhtml">
<url>
<loc>http://www.example.com/page-1/</loc>
<xhtml:link
  rel="alternate"
  media="only screen and (max-width: 640px)"
  href="http://m.example.com/page-1"/>
</url>
</urlset>

<?xml version="1.0" encoding="UTF-8"?>
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"
  xmlns:mobile="http://www.google.com/schemas/sitemap-mobile/1.0">
<url>
  <loc>http://mobile.example.com/article100.html</loc>
  <mobile:mobile/>
</url>
</urlset>
```

Εικόνα 5.9: Sitemap για τη σελίδα στον υπολογιστή και στη κινητή συσκευή αντίστοιχα[29].

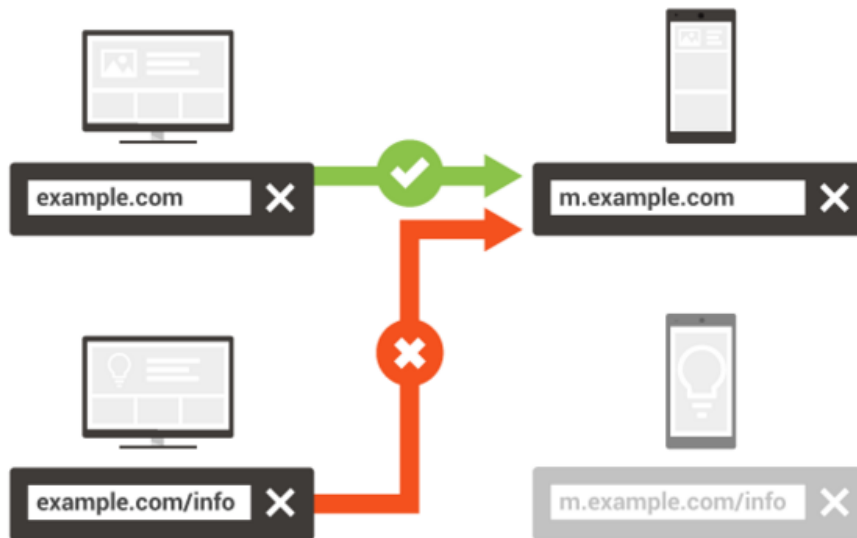
Εξήγηση των χαρακτηριστικών(attributes) της ετικέτας(tag) του συνδέσμου για τη σελίδα στους επιτραπέζιους υπολογιστές:

- Το χαρακτηριστικό (attribute) rel="alternate" χρησιμοποιείται ώστε να γνωστοποιήσει ότι υπάρχει εναλλακτικός σύνδεσμος για τη σελίδα που εμφανίζεται στους υπολογιστές.
- Η τιμή για το χαρακτηριστικό(attribute) της CSS περιγράφει στη google πότε πρέπει να χρησιμοποιηθεί ο εναλλακτικός σύνδεσμος (url).
- Το χαρακτηριστικό(attribute) href ορίζει τη τοποθεσία στην οποία βρίσκεται ο εναλλακτικός σύνδεσμος(url). Που στη προκειμένη περίπτωση ως εναλλακτικό σύνδεσμος ώστε να χρησιμοποιηθεί από τις κινητές συσκευές έχουμε ορίσει την m.example.com/page-1.

Αυτή η αμφίδρομη στάση βοηθάει το ρομπότ να ανακαλύψει το περιεχόμενο των σελίδων καθώς και τους αλγορίθμους να καταλάβουν τη σχέση μεταξύ των σελίδων των υπολογιστών και των κινητών συσκευών αντίστοιχα. Η χρησιμοποίηση ξεχωριστών συνδέσμων(url) για το σερβίρισμα του ίδιου περιεχομένου σε διάφορες μορφές βοηθάει ώστε να αντιμετωπίζονται ως μία οντότητα αντί για δύο. Αν το ρομπότ δεν είναι σε θέση να καταλάβει την αντιστοιχία μεταξύ των και τις αντιμετωπίσει ως 2 οντότητες μπορεί η κατάταξη

τους να διαφέρει. Επιπρόσθετα τα πιο συνήθως λάθη που μπορεί να συμβούν είναι :

- Κατά τη χρησιμοποίηση του `rel="alternate"` και του `rel="canonical"` πρέπει να διατηρείται μια σχέση ένα προς ένα(1-1) μεταξύ σύνδεσης της σελίδας που αντιστοιχεί στον υπολογιστή και στη κινητή συσκευή.
- Απαιτείται ιδιαίτερος έλεγχος στις ανακατευθύνσεις ώστε να έχουν τους σωστούς συνδέσμους.



Εικόνα 5.10: Σχέση 1 προς 1 μεταξύ σελίδας που αντιστοιχεί στον υπολογιστή και στη κινητή συσκευή με σωστές ανακατευθύνσεις[28].

Η δημιουργία αυτής της σχέσης μπορεί να γίνει με δύο είδους τεχνικές:

HTTP ανακατευθύνσεις: Ο συγκεκριμένος τρόπος αποτελεί συνηθισμένη τεχνική για την ανακατεύθυνση των διευθύνσεων(url). Οι ανακατευθύνσεις βασίζονται στο http header και είναι σημαντικό να ορίζονται στη σελίδα με το `rel="alternate"` ή στο sitemap. Δεν έχει σημασία αν ο server ανακατευθύνει με HTTP 301 ή 302 κωδικό κατάστασης (status code), όμως η χρήση του 302 συστήνεται όπου είναι δυνατόν.

Javascript ανακατευθύνσεις: Εάν η ανακατεύθυνση είναι δύσκολο να εφαρμοστεί με HTTP, μπορεί να χρησιμοποιηθεί η javascript. Αν επιλεγεί η συγκεκριμένη μέθοδος, θα πρέπει να λάβουμε υπόψη κάποια καθυστέρηση από τη πλευρά του χρήστη. Αυτό θα συμβεί γιατί θα χρειαστεί να γίνει πρώτα μια λήψη

Τεχνικές, μέθοδοι και κανόνες για τη βελτιστοποίηση ανάπτυξης ιστοσελίδων

της σελίδας πριν εκτελεστεί η javascript ώστε να κάνει τη σωστή ανακατεύθυνση. Υπάρχουν πολλές προσεγγίσεις για την υλοποίηση μιας ανακατεύθυνσης μέσω javascript. Για παράδειγμα μπορεί να υλοποιηθεί με την συνάρτηση(function) `matchMedia()`[28].

6 SEM- Marketing στις MA

Μέχρι τώρα είδαμε πως η αναζήτηση πληροφοριών και η ανάδειξη τους γινόταν με βάση κάποιους συγκεκριμένους κανόνες που ορίζουν οι MA. Η μέθοδος όμως που οδηγεί στη τελική κατάταξη των ιστοσελίδων ονομάζεται SEM(Search Engine Marketing). Στο συγκεκριμένο κεφάλαιο θα δούμε τι είναι το SEM και πως αυτό σχετίζεται με το SEO.

6.1 Ορισμός SEM

Το μάρκετινγκ των MA(SEM) στοχεύει στη προώθηση ιστοσελίδων στο διαδίκτυο, στη καλύτερη προβολή αυτής, έχοντας ως απώτερο σκοπό την εμφάνιση στα υψηλότερα αποτελέσματα των MA. Το SEM χωρίζεται σε δύο κλάδους. Ο πρώτος είναι η βελτιστοποίηση των ιστοσελίδων για τις MA ή SEO και ο δεύτερος μέσω διαφημίσεων που ονομάζονται pay-per click(PPC).



Εικόνα 6.1: Ορισμός SEM.

Σήμερα οι MA εμφανίζουν οργανικά και μη αποτελέσματα στην ίδια αναζήτηση που κάνει ο χρήστης. Οργανικά ορίζονται τα αποτελέσματα τα οποία εμφανίζονται στα ερωτήματα του χρήστη και δε συμπεριλαμβάνονται σε αυτά διαφημίσεις. Το γεγονός ότι μια ιστοσελίδα πληρώνει για διαφήμιση δε τη κατατάσσει αυτόματα και στα πρώτα αποτελέσματα των MA.

Γυναικεία αθλητικά ρούχα - Ntynomai.gr

www.ntynomai.gr > Γυναικεία > Ρούχα ▼

Γυναικεία αθλητικά ρούχα από τα μεγαλύτερα Ελληνικά και Ευρωπαϊκά ηλεκτρονικά καταστήματα.

Αθλητικά Ρούχα

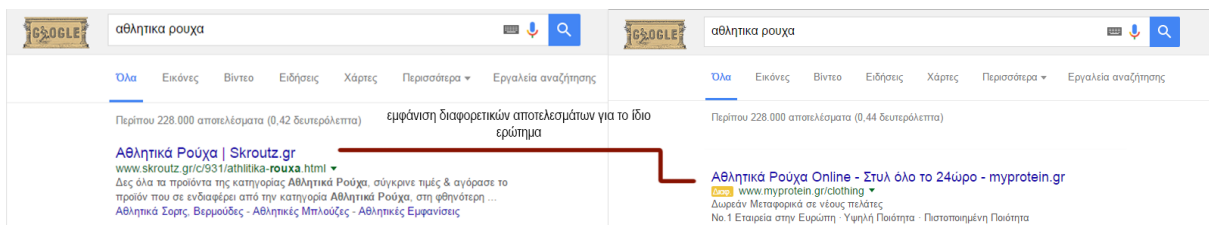
Διαφ. www.esports4u.gr/ ▼

Όμορφα επώνυμα αθλητικά ρούχα σε μεγάλη ποικιλία & χαμηλές τιμές!

Vision Street Wear - Raw craft uk φουτερ - Ecko Unltd - Φουτερ No Fear

Εικόνα 6.2: Διαφημίσεις εναντίον seo.

Επίσης πρέπει να τονιστεί πως για τους χρήστες που χρησιμοποιούν κάποια επέκταση(extension) η οποία απαγορεύει στο περιηγητή(browser) την εμφάνιση διαφημίσεων, αυτοί δε θα μπορούν να βλέπουν τις σελίδες που έχουν επιλέξει το συγκεκριμένο τρόπο προώθησης.



Εικόνα 6.3: Τα extension επηρεάζουν το τρόπο εμφάνισης αποτελεσμάτων.

6.1.1 PPC

Η διαφήμιση με χρέωση ανά κλικ (Pay Per Click/PPC) είναι η μέθοδος που περιλαμβάνει την τιμή προσφοράς που θα χρεωθεί ο διαφημιζόμενος, όταν ο χρήστης των MA επιλέξει (κάνει κλικ) στην διαφήμιση του. Μία από τις δημοφιλέστερες πλατφόρμες διαφήμισης, είναι το Google Adwords, που εμφανίζει την διαφημιστική χορηγία στην περιοχή της λίστας των αποτελεσμάτων αναζήτησης της Google[42].

Βασικοί όροι PPC

Αναλογία κλικ/εμφανίσεις(CTR): Η αναλογία των κλικ ανά εμφανίσεις ή αλλιώς ο όρος CTR (Click Through Rate), χρησιμοποιείται για την μέτρηση των αποτελεσμάτων της διαφήμισης PPC (Pay Per Click). Η αναλογία των κλικ /

εμφανίσεων(CTR) αναφέρεται στον αριθμό των επισκεπτών που επέλεξαν μία συγκεκριμένη διαφήμιση και βασίζεται στις εμφανίσεις και στην δυναμική των αναζητήσεων της. Επιπλέον, το CTR μπορεί να μετρήσει την αποτελεσματικότητα των στοιχείων του μάρκετινγκ που ακολουθούνται σε μία ιστοσελίδα.

Προσφορά κόστους ανά χίλιες εμφανίσεις(CPM): Η προσφορά κόστους ανά χίλιες εμφανίσεις, είναι ένας τρόπος κοστολόγησης των διαφημίσεων μέσω διαδικτύου, διαφορετικός από την χρέωση ανά κλικ (PPC). Με το CPM το κόστος της διαφήμισης ορίζεται στις 1000 εμφανίσεις ή επισκέψεις ανεξαιρέτως πόσα κλικ (επιλογές) θα συγκεντρώσει[42].

Κόστος Ανά Κλικ (Cost Per Click, CPC): Το ποσό που καταβάλλεται για κάθε κλικ σε μια διαφήμιση η οποία δημιουργείτε μέσω μιας πλατφόρμας πληρωμένων διαφημίσεων όπως το Google AdWords, το Yahoo Search Marketing και το Microsoft adCenter[47].

Conversion rate: Είναι ένα ποσοστό το οποίο υπολογίζεται από τον αριθμό των χρηστών που πραγματοποιούν μια συγκεκριμένη δραστηριότητα π.χ. αγοράζουν ένα προϊόν μας, προς τον συνολικό αριθμό των χρηστών που προσεγγίζουμε[43].

6.2 Διαφορές SEO και PPC και κριτήρια επιλογής

Βασικός σκοπός και των 2 μεθόδων αποτελεί η αύξηση της επισκεψιμότητας(traffic) της ιστοσελίδας μας. Η διαφορά τους είναι στο κόστος και στο χρόνο που απαιτείται για την υλοποίηση του καθενός. Πιο συγκεκριμένα το κόστος για το SEO είναι εφάπαξ ενώ στη περίπτωση του PPC υπάρχει χρέωση για κάθε φορά που κάποιος εισέρχεται στη ιστοσελίδα μας μέσω της διαφήμισης. Όσον αφορά το παράγοντα του χρόνου αν επιλέξουμε τη τεχνική του seo θα χρειαστούμε περισσότερο καθώς απαιτείται κάποια προεργασία στην ιστοσελίδα μας για την εφαρμογή των white-hat κανόνων και έπειτα ευρετηρίασης(indexing) της ιστοσελίδας μας από τις MA. Και οι δύο τρόποι μπορούν να μας φέρουν καλά αποτελέσματα. Το πιο σημαντικό κριτήριο το οποίο πρέπει να σκεφτούμε για την

Τεχνικές, μέθοδοι και κανόνες για τη βελτιστοποίηση ανάπτυξης ιστοσελίδων

επιλογή του τρόπου προώθησης της ιστοσελίδας μας είναι το οικονομικό κομμάτι, καθώς και πόσο γρήγορα μπορεί να γίνει η απόσβεση των χρημάτων που θα διαθέσουμε. Τέλος, προτείνεται η εφαρμογή και των δύο μεθόδων σε περίπτωση που μπορεί να διατεθεί κάποιο ποσό για διαφήμιση.

7 Βιβλιογραφία

- [1] https://el.wikipedia.org/wiki/%CE%94%CE%B9%CE%B1%CE%B4%CE%AF%CE%BA%CF%84%CF%85%CE%BF#.CE.A4.CE.BF_.CE.94.CE.B9.CE.B1.CE.B4.CE.AF.CE.BA.CF.84.CF.85.CE.BF_.CE.BA.CE.B1.CE.B9_.CE.B7_.CE.95.CF.80.CE.B9.CE.BA.CE.BF.CE.B9.CE.BD.CF.89.CE.BD.CE.AF.CE.B1
- [2] https://el.wikipedia.org/wiki/%CE%9C%CE%B7%CF%87%CE%B1%CE%BD%CE%AE_%CE%B1%CE%BD%CE%B1%CE%B6%CE%AE%CF%84%CE%B7%CF%83%CE%B7%CF%82
- [3] <http://osarena.net/faqs/search-engines-history-ke-alla.html>
- [4] <https://el.wikipedia.org/wiki/Yahoo!>
- [5] <https://en.wikipedia.org/wiki/MSN>
- [6] <https://en.wikipedia.org/wiki/Yandex>
- [7] <http://www.seoingreece.org/%CE%BC%CE%B7%CF%87%CE%B1%CE%BD%CE%AD%CF%82-%CE%B1%CE%BD%CE%B1%CE%B6%CE%AE%CF%84%CE%B7%CF%83%CE%B7%CF%82/>
- [8] <https://tools.ietf.org/html/rfc3986>
- [9] https://en.wikipedia.org/wiki/URL_normalization
- [10] https://en.wikipedia.org/wiki/Search_engine_indexing#Index_data_structures
- [11] http://pacific.jour.auth.gr/totsidou/Search_Engines_in_the_WWW.htm
- [12] Michael A. and Salter B. (2008) Marketing through Search Optimization: How people search and how to be found on the web. 2nd Edition. USA: Elsevier
- [13] <https://searchenginewatch.com/sew/news/2065173/how-search-engines-work>
- [14] Michie J. (2006) Street Smart Internet Market. Canada: Performance Marketing Group
- [15] https://en.wikipedia.org/wiki/Metasearch_engine
- [16] https://el.wikipedia.org/wiki/Google#cite_note-2
- [17] <https://en.wikipedia.org/wiki/Google>
- [18] <http://infolab.stanford.edu/~backrub/google.html>
- [19] <https://en.wikipedia.org/wiki/PageRank>
- [20] http://news.oneseocompany.com/2012/07/15/why-google-is-the-leading-search-engine-website_20120715740.html
- [21] <http://www.businessinsider.com/accurate-contrarian-theories-2011-5>

[22] <https://en.wikipedia.org/wiki/Bing>

[23] <http://lifehacker.com/the-best-things-bing-does-better-than-google-1483983630>

[24] <https://www.clickz.com/clickz/column/2288982/search-in-china-how-baidu-is-different-from-google>

[25] <https://en.wikipedia.org/wiki/Baidu>

[26] <https://developers.google.com/webmasters/mobile-sites/mobile-seo/responsive-design>

[27] <https://developers.google.com/webmasters/mobile-sites/mobile-seo/dynamic-serving>

[28] <https://developers.google.com/webmasters/mobile-sites/mobile-seo/separate-urls#tldr>

[29] <https://www.branded3.com/blog/implementing-mobile-site-annotations/>

[30] <http://www.seoingreece.org/panda-4-0/>

[31] <http://www.seoingreece.org/search-engine-optimization/>

[32] <http://www.seoingreece.org/black-hat-seo/>

[33] <http://digitalmarketingphilippines.com/the-pros-and-cons-of-using-seo-for-your-business/>

[34] <http://www.seoingreece.org/white-hat-seo/>

[35] <http://static.googleusercontent.com/media/www.google.com/el//webmasters/docs/search-engine-optimization-starter-guide.pdf>

[36] <https://moz.com/learn/seo/internal-link>

[37] <http://www.seoingreece.org/%CF%80%CF%81%CE%BF%CF%8E%CE%B8%CE%B7%CF%83%CE%B7-%CE%B9%CF%83%CF%84%CE%BF%CF%83%CE%B5%CE%BB%CE%AF%CE%B4%CF%89%CE%BD-backlinks/>

[38] <http://www.seoingreece.org/social-media-seo/>

[39] https://en.wikipedia.org/wiki/Google_Search_Console

[40] <http://www.seoingreece.org/seo-tools/>

[41] https://d1avok0lzl2w.cloudfront.net/img_uploads/web-app-1.gif

[42] <http://business.vrisko.gr/blog/2016/ianouarios/web-marketing-terms>

[43] <http://www.seoingreece.org/%CE%BF%CF%81%CE%BF%CE%BB%CE%BF%CE%B3%CE%AF%CE%B1-seo/>

[44] <https://support.google.com/webmasters/answer/35769?hl=en>

[45] <https://cs.nyu.edu/courses/fall02/G22.3033-008/lec10.html>

[46] <http://www.au.af.mil/au/>, (2014). 42. [online] Available at: [http://Intelligence Exploitation of the internet](http://IntelligenceExploitationoftheinternet) [Accessed 31 Oct. 2014]

[47] <http://www.goseo.gr/kentro-voithias/seo-sem-lexilogio-oron/#C>