



ΑΛΕΞΑΝΔΡΕΙΟ Τ.Ε.Ι. ΘΕΣΣΑΛΟΝΙΚΗΣ  
ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΚΩΝ ΕΦΑΡΜΟΓΩΝ  
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ



Πτυχιακή Εργασία

# Βελτιστοποίηση Ιστοσελίδων για Μηχανές Αναζήτησης



Του φοιτητή  
Σελημίδη Δημητρίου  
Αρ. Μητρώου: 05/2762

Επιβλέπων Καθηγητής  
Φώτης Κόκορας

Θεσσαλονίκη 2010



# Πρόλογος

Λόγω της μη λεπτομερούς κοινοποίησης του μηχανισμού ταξινόμησης των αποτελεσμάτων από τις μηχανές αναζήτησης, γίνονται πολλές εικασίες σχετικά με το ποια στοιχεία μιας ιστοσελίδας επηρεάζουν την κατάταξή της στα αποτελέσματα αναζήτησης με βάση κάποια λέξη-κλειδί. Το όλο θέμα απαντάται στη βιβλιογραφία με τον όρο *search engine optimization* όπου προφανώς το *optimization* (βελτιστοποίηση) αφορά στην υπό εξέταση κάθε φορά ιστοσελίδα.

Στόχος της πτυχιακής είναι η καταγραφή των στοιχείων που για κάποιους λόγους πιστεύεται ότι επηρεάζουν της παραπάνω κατάταξη, καθώς και η ανάπτυξη εφαρμογής που θα παρέχει πληροφόρηση για τη συμμόρφωση ή όχι κάποιας ιστοσελίδας με τους σχετικούς κανόνες που υπονοούνται.

Η παρούσα πτυχιακή εργασία εκπονήθηκε στο πλαίσιο του σχετικού μαθήματος του Τμήματος Πληροφορικής, της σχολής Τεχνολογικών Εφαρμογών του Αλεξάνδρειου Τεχνολογικού Ιδρύματος Θεσσαλονίκης, υπό την επίβλεψη του κ. Φώτη Κόκκορα, επιστημονικού συνεργάτη του ΤΕΙ.

Ιδιαίτερες ευχαριστίες απευθύνονται στον επιβλέποντα καθηγητή για την καθοδήγησή του σε όλη την διάρκεια της πτυχιακής, την παροχή βιβλιογραφικών και όχι μόνο πηγών και την αμεσότητα και πληρότητα στην παροχή διευκρινήσεων σε όσες απορίες προέκυπταν. Ευχαριστίες απευθύνονται επίσης στην εταιρία *New Media - Τεχνολογίες Νέων Μέσων ΑΕ* και στο προσωπικό της για την ανταλλαγή απόψεων, την βοήθεια, τα σεμινάρια αλλά και την μετάβαση από την θεωρία στην πράξη καθ' όλη την διάρκεια της εκπόνησης της πρακτικής άσκησης στην εταιρεία.

Σελημίδης Δημήτριος



# Περίληψη

Η βελτιστοποίηση μιας ιστοσελίδας ώστε αυτή να ευνοείται στην κατάταξη των αποτελεσμάτων αναζήτησης είναι ένα σχετικά νέο θέμα στον χώρο των διαδικτυακών εφαρμογών που όμως έχει λάβει μεγάλη έκταση και απασχολεί πλέον έντονα όσους ασχολούνται με ιστοσελίδες, είτε είναι προγραμματιστές ιστοτόπων είτε ιδιοκτήτες.

Η ανάγκη της βελτιστοποίησης των ιστοσελίδων πηγάζει από τον στόχο που θέτει κάθε ιστοσελίδα να βρίσκεται στις υψηλότερες θέσεις στα αποτελέσματα μιας μηχανής αναζήτησης. Με βάση παραμέτρους που είναι γνωστό ότι λαμβάνονται υπόψη από τις μηχανές αναζήτησης καθώς και άλλες που εικάζεται ότι λαμβάνονται υπόψη, έχει δημιουργηθεί ένα σύνολο από κανόνες δόμησης ιστοσελίδων που αποσκοπούν στο να κριθεί ευνοϊκότερα η εκάστοτε ιστοσελίδα σε δεδομένη αναζήτηση.

Βέβαια, με τον καιρό, η ισχύς κάποιων χαρακτηριστικών επιβεβαιώνεται, ενώ κάποιων άλλων όχι. Επίσης μπορεί κάτι που ίσχυε παλαιότερα να πάψει να υφίσταται. Ως αποτέλεσμα η βελτιστοποίηση ιστοσελίδων αποτελεί κατά ένα βαθμό κινούμενο στόχο και απαιτείται συνεχής ενημέρωση και παρακολούθηση των εξελίξεων του τομέα. Όσες τεχνικές αναφέρονται στην πτυχιακή εργασία είναι, στην παρούσα χρονική στιγμή, επιβεβαιωμένες. Κάποιες όμως από αυτές μπορεί να τροποποιηθούν ή ακόμη και να καταργηθούν στο μέλλον.

Οι πρακτικές που ακολουθούνται χωρίζονται σε καλές και κακές ανάλογα με το πώς εκλαμβάνονται μακροπρόθεσμα από μηχανές αναζήτησης. Οι λεγόμενες «κακές πρακτικές» βασίζονται σε κατάχρηση κανόνων για βραχυπρόθεσμα οφέλη που όμως δημιουργούν το ρίσκο της δυσμενούς αντιμετώπισης (black list) ενός ιστοχώρου από τις μηχανές αναζήτησης, αν γίνουν αντιληπτά από αυτές. Τάση της εποχής είναι η μίσθωση επαγγελματιών αλλά και χρήση ειδικών λογισμικών για την ανάλυση των στοιχείων που επηρεάζουν την κατάταξη μιας ιστοσελίδας.

Για τον σκοπό αυτό και στο πλαίσιο της πτυχιακής, έγινε αναζήτηση σε πολλές πηγές που μιλούν για βελτιστοποίηση και μετά από διασταύρωση των

απόψεων αλλά και επιβεβαίωση της ισχύος των τεχνικών είτε από ειδήμονες είτε πειραματικά, καταγράφηκαν και αναλύθηκαν αυτές οι παράμετροι. Επιπλέον δημιουργήθηκε και σχετική εφαρμογή που αναλύει την δομή δοθείσας ιστοσελίδας σε σχέση με τις λέξεις-κλειδιά που αυτή περιέχει και τη βαθμολογεί με βάση το βαθμό συμμόρφωσής της με τους σχετικούς κανόνες.

# Abstract

Web page optimization for better ranking in search engine page results, SERP, is a relatively new topic on the World Wide Web which has extensively grown and affects those who involve in websites such as website owners and developers.

The need of web site engine optimization comes from the goals a website sets to be ranked in higher position in SERP. Based on the parameters which are known that search engines take into consideration, there is a set of website building rules to determine the most favorable site to each given search.

Over time, the validity of some rules is confirmed, while some others do not. Some rules that worked previously cease to exist. As a result, optimizing web sites is a moving target and requires constant searching and reading information about SEO. Those techniques mentioned in this thesis, at this time, are confirmed. But some of them may be amended or even abolished in the future.

These practices are divided into good and bad depending on how search engine judge. The so-called «bad practices» are based on the abuse of rules for short-term benefits, but has the risk of adverse response (black list) of a website from search engines, if the second one finds the abuse. Today's trend is hiring professionals and use of specific software to analyze the elements affecting the ranking of a website

For this thesis, a research on many sources that talk about search engine optimization has been made and after crossing information and confirmation of these techniques by experts or by experimenting, were recorded and analyzed. Moreover, an application has been created that analyzes the structure of the given site based on keywords and gives a final score.





# Περιεχόμενα

Πρόλογος .....	1
Περίληψη.....	3
Abstract.....	5
Περιεχόμενα .....	7
1 Εισαγωγή.....	11
2 Βελτιστοποίηση Σελίδων για Μηχανές Αναζήτησης.....	19
2.1 Βασικές Έννοιες & Ιστορική Αναδρομή .....	19
2.1.1 Λέξεις-Κλειδιά .....	19
2.1.2 Page Rank.....	20
2.1.3 Βελτιστοποίηση & Μάρκετινγκ (SEM).....	22
2.1.4 PPC ή Βελτιστοποίηση .....	24
2.1.5 Καλή και Κακόβουλη Βελτιστοποίηση.....	26
2.1.6 Φιλικές διευθύνσεις.....	27
2.1.7 Βελτιστοποίηση Σελίδων για Κινητά Τηλέφωνα .....	28
2.1.8 Βελτιστοποίηση Σελίδων και Ελληνική Γλώσσα .....	30
2.1.9 Ιστορική Αναδρομή.....	32
2.1.10 Βελτιστοποίηση Σελίδων και Κοινωνικά Δίκτυα .....	35
2.1.11 Βελτιστοποίηση Δυναμικών ιστοσελίδων.....	36
2.2 Στοιχεία και Τεχνικές .....	39
2.2.1 Domain .....	39
2.2.2 Τίτλος Σελίδας .....	41
2.2.3 Περιγραφή Ιστοσελίδας .....	42
2.2.4 Ετικέτες Επικεφαλίδων .....	43
2.2.5 Περίληψη Σελίδας.....	44

2.2.6	Έντονες/Πλάγιες Λέξεις.....	45
2.2.7	Υπερσυνδέσεις .....	46
2.2.8	Alt tags.....	48
2.2.9	Πυκνότητα Λέξεων-Κλειδιών .....	50
2.2.10	Πλοήγηση Σελίδας .....	51
2.2.11	Καθαρός Κώδικας.....	52
2.2.12	Rewrite Engine.....	53
2.3	Κακόβουλη Βελτιστοποίηση .....	56
2.3.1	Redirect Hijacking .....	56
2.3.2	Page Swapping (Cloaking) .....	58
2.3.3	Doorway & Σελίδες Πληροφοριών.....	60
2.3.4	Κρυφό Κείμενο.....	60
2.3.5	Άλλες "Μαύρες" Τεχνικές .....	61
3	Λύσεις Βελτιστοποίησης.....	63
3.1	Βελτιστοποίηση κατά τη Σχεδίαση.....	63
3.2	Βελτιστοποίηση εκ των Υστέρων.....	64
3.3	Ατομική Εργασία ή Μίσθωση Επαγγελματία .....	65
4	Εργαλεία Βελτιστοποίησης.....	67
4.1	Αναλυτές Βελτιστοποίησης.....	67
4.2	Έρευνα και υπόδειξη λέξεων-κλειδιών .....	69
4.3	Ανάλυση πυκνότητας λέξεων-κλειδιών .....	70
4.4	Υπολογιστής Page Rank .....	71
4.5	Δημοτικότητα συνδέσμου .....	71
4.6	Προσθήκη σε Καταλόγους.....	72
4.7	Εντολές αναζήτησης της Google .....	73
4.8	Seo-Browser.com.....	75
5	SEO Analyzer .....	77

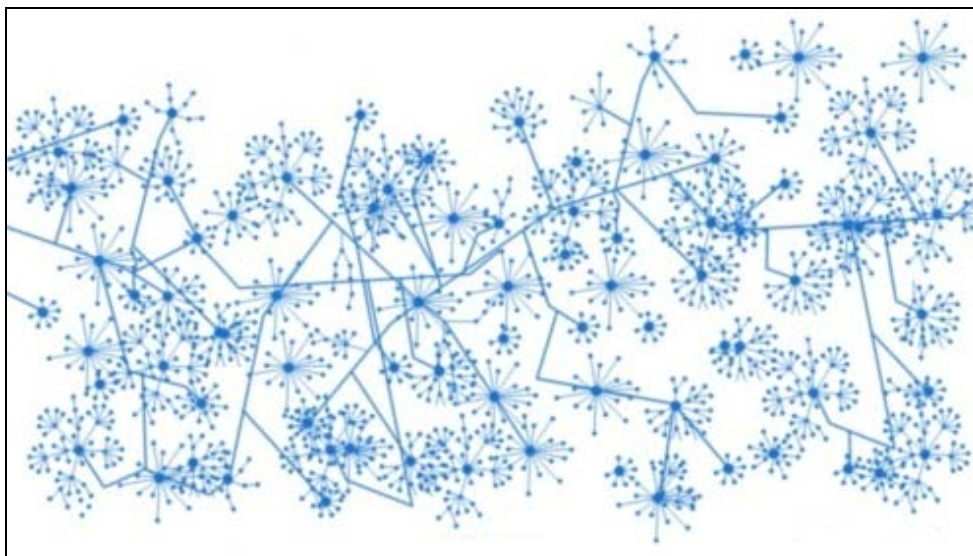
5.1	Γενική Χρήση .....	77
5.2	Τεχνική Περιγραφή.....	80
5.3	Τεχνικές Δυσκολίες .....	81
6	Συμπεράσματα .....	83
	Αναφορές .....	87
	Βιβλιογραφία .....	89
	Παράρτημα.....	91



# 1 Εισαγωγή

Το διαδίκτυο αποτελεί ένα από τα μεγαλύτερα επιτεύγματα του ανθρώπου. Η παγκόσμια ένωση των ηλεκτρονικών υπολογιστών κάτω από ένα μεγάλο δίκτυο και ο χρόνος μετάδοσης πληροφοριών, που κυμαίνεται σε κλάσματα δευτερολέπτων, είναι από τα στοιχεία που το κάνουν μοναδικό εργαλείο αλλά και βάση για την ανάπτυξη πολλών άλλων εργαλείων και υπηρεσιών. Η εξέλιξη της τεχνολογίας επηρεάστηκε πολύ από την ταχύτητα εξάπλωσης του και υιοθετείτε σε όλο και περισσότερα προϊόντα.

Μια από τις σημαντικότερες λειτουργίες του διαδικτύου είναι ο Παγκόσμιος Ιστός. Ο Παγκόσμιος Ιστός αποτελείται ένα απέραντο δίκτυο από ιστοσελίδες, αλληλένδετες μεταξύ τους, που μεταδίδουν πληροφορίες. Οι ιστοσελίδες είναι σελίδες υπερκειμένου (hypertext), έγγραφα προσβάσιμα μέσω διαδικτύου. Μια ιστοσελίδα περιέχει κείμενο ή πολυμέσα, όπως βίντεο, κινούμενες εικόνες και ήχους. Το κύριο χαρακτηριστικό των ιστοσελίδων είναι ότι συνδέονται με άλλες εσωτερικές σελίδες ή άλλες σελίδες ιστοχώρων. Σχηματικά, η ένωση των ιστοσελίδων θα μπορούσε να χαρακτηριστεί ως ένας ιστός αράχνης, όπως φαίνεται στην Εικόνα 1.1: Δείγμα αλληλένδετων ιστοσελίδων, γι αυτό και υιοθετήθηκε ως εικόνα και όνομα η φράση «Παγκόσμιος ιστός»

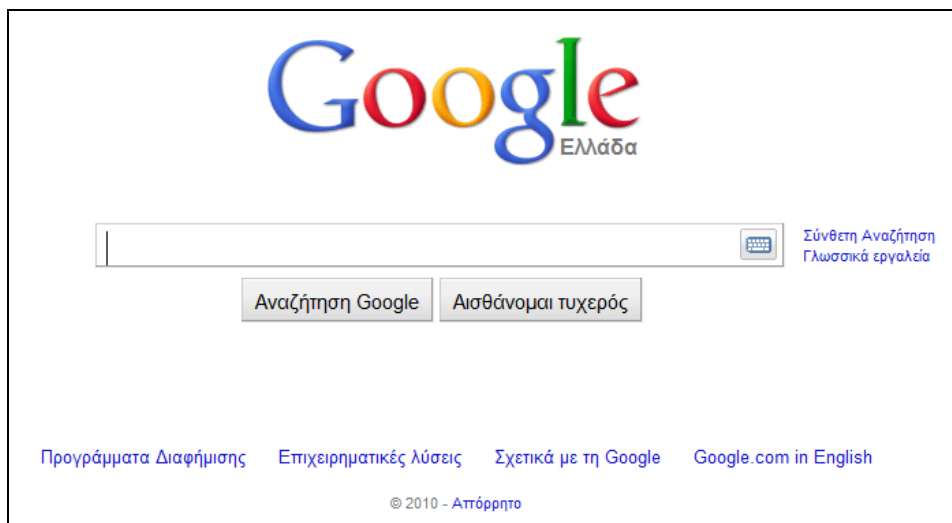


Εικόνα 1.1: Δείγμα αλληλένδετων ιστοσελίδων

Καθημερινά, ο αριθμός των ιστοσελίδων αυξάνεται εκθετικά. Όλο και περισσότερες επιχειρήσεις στρέφονται προς το ηλεκτρονικό εμπόριο αλλά και χρήστες του διαδικτύου θέλουν να μοιραστούν τις γνώσεις τους, τις απόψεις τους, ακόμη και να συνομιλήσουν με άλλους σε διάφορα θέματα. Η θεαματική αύξηση των ιστοσελίδων έφερε την ανάγκη ύπαρξης εργαλείων ικανά να ευρετηριάσουν ιστοσελίδες από όλο το διαδίκτυο και να επιτρέπουν στον χρήστη να δίνει ορισμένες λέξεις ή φράσεις και να αναζητά τα δεδομένα μέσα στα ευρετήρια. Τα αποτελέσματα της αναζήτησης θα πρέπει να είναι έγκυρα με τα δεδομένα του χρήστη.

Για τον σκοπό αυτό δημιουργήθηκαν οι μηχανές αναζήτησης, εργαλεία κατάλληλα σχεδιασμένα για την αναζήτηση πληροφοριών στο Διαδίκτυο. Είναι ίσως οι πιο υψηλά, σε επισκέπτες, ιστοσελίδες μιας που προσφέρουν μεγάλη διευκόλυνση στην εύρεση πληροφοριών. Η χρήση των μηχανών αναζήτησης είναι απλή και δωρεάν.

Όπως φαίνεται και στην Εικόνα 1.2 ο χρήστης συμπληρώνει σε ένα πεδίο ορισμένες λέξεις ή φράσεις σχετικά με τις πληροφορίες που αναζητά και η μηχανή αναζήτησης του επιστρέφει μια σελίδα αποτελεσμάτων. Τα αποτελέσματα είναι μια λίστα από ιστοσελίδες, ταξινομημένη ανάλογα με το ποσοστό σχετικότητας της κάθε σελίδας ως προς τους όρους αναζήτησης. Οι μηχανές αναζήτησης όμως δεν περιορίζονται στην αναζήτηση κειμένου. Σύγχρονες μηχανές αναζήτησης παρέχουν την δυνατότητα εύρεσης εικόνων, βίντεο καθώς και άλλα είδη αρχείων.

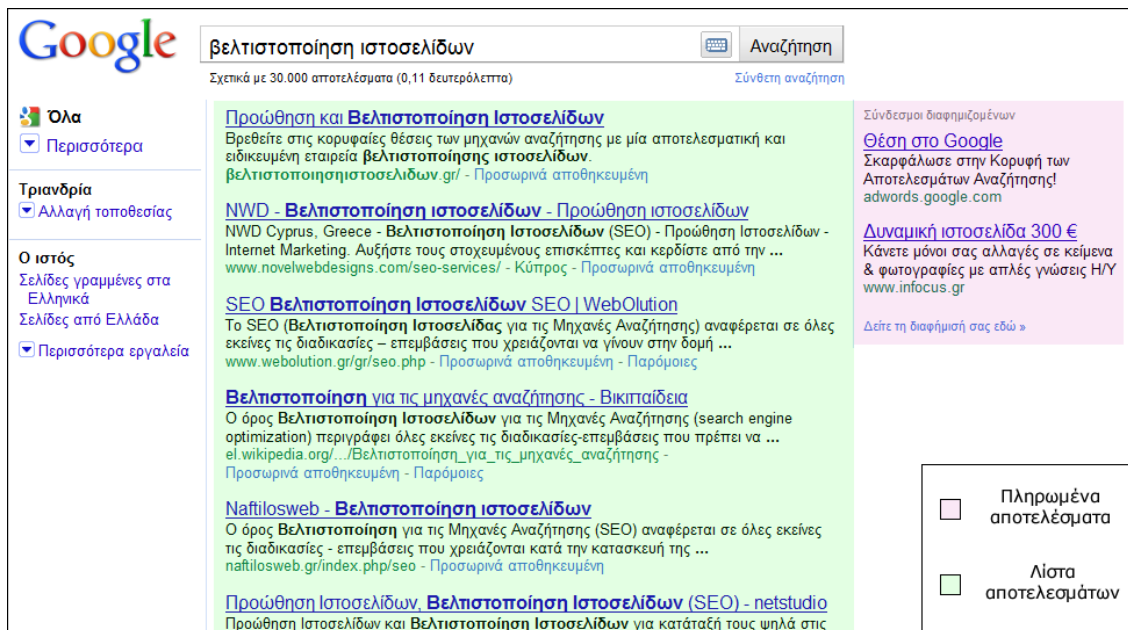


Εικόνα 1.2: Μηχανή αναζήτησης Google

Οι μηχανές αναζήτησης λειτουργούν βάσει τις πληροφορίες που συλλέγουν και στηρίζουν την λειτουργία τους στην βασική ουσία του παγκόσμιου ιστού. Κάθε μηχανή αναζήτησης έχει έναν Web Crawler (ή spider), έναν αυτόματο φυλλομετρητή που επισκέπτεται κάθε σύνδεσμο που εντοπίζει σε μια σελίδα. Ο ρόλος του είναι να συλλέγει πληροφορίες από κάθε σελίδα που επισκέπτεται, να τις αναλύει και να τις αποθηκεύει σε ένα ευρετήριο με την χρήση κάποιων κριτηρίων. Πιο προηγμένες μηχανές αναζήτησης αποθηκεύουν και μέρος της σελίδας ως στιγμιότυπο για διάφορες χρήσεις, όπως η προεπισκόπηση της από τον διακομιστή της μηχανής αναζήτησης όταν η σελίδα δεν λειτουργεί προσωρινά.

Όταν ο χρήστης εισάγει κάποια λέξη ή φράση σχετικά με το αντικείμενο που αναζητά, ο αλγόριθμος αναζήτησης ψάχνει μέσα στο ευρετήριο για ιστοσελίδες που είναι σχετικές με τους όρους αναζήτησης και επιστρέφει μια λίστα αποτελεσμάτων από τις πιο σχετικές ιστοσελίδες. Εκτός από λέξεις ή φράσεις ο χρήστης μπορεί να χρησιμοποιήσει ορισμένες εντολές όπως AND, OR ή ειδικές εντολές-φίλτρα για πιο παραμετροποιημένη ή εξειδικευμένη αναζήτηση.

Τα αποτελέσματα αναζήτησης είναι διαμορφωμένα ανάλογα την σχετικότητα κάθε ιστοσελίδας με τους όρους αναζήτησης αλλά και την βαρύτητα που τις δίνει η έκαστη μηχανή αναζήτησης. Όπως παρατηρείται στην Εικόνα 1.3: Σελίδα αποτελεσμάτων του όρου "βελτιστοποίηση ιστοσελίδων" συναντάται συχνά, η τοποθέτηση πληρωμένων αποτελεσμάτων, γνωστά και ως σπόνσορες ή διαφημιστές στην κορυφή ή στο πλαϊνό μέρος της σελίδας των αποτελεσμάτων. Η περιοχή εμφάνισης τους είναι ευδιάκριτη αλλά γνωστοποιείται ότι πρόκειται για πληρωμένες υπερσυνδέσεις. Αυτό είναι το κύριο έσοδο των μηχανών αναζήτησης.



Εικόνα 1.3: Σελίδα αποτελεσμάτων του όρου "βελτιστοποίηση ιστοσελίδων"

Η σειρά εμφάνισης των ιστοσελίδων στην λίστα αποτελεσμάτων διαφέρει από μια μηχανή αναζήτησης με μια άλλη διότι κάθε μια χρησιμοποιεί διαφορετικό αλγόριθμο ευρετηρίασης και κατάταξης αποτελεσμάτων. Για παράδειγμα, αναζητώντας την λέξη «game» στην Google ([www.google.com](http://www.google.com)) και στην Altavista ([www.altavista.com](http://www.altavista.com)), εμφανίζεται η σελίδα της Miniclip και της Wikipedia αντίστοιχα – αποτελέσματα αναζήτησης 20/10/09.

Η Google είναι η πρώτη, σε προτίμηση χρηστών, μηχανή αναζήτησης του διαδικτύου. Οι υπηρεσίες που προσφέρει απέχουν πολύ από τις κλασσικές μηχανές αναζήτησης και μονοπωλεί το ενδιαφέρον των χρηστών. Παρόλο που ή Yahoo! ήταν από τις πρώτες μηχανές αναζήτησης, η Google βρίσκεται στην κορυφή. Για τον τρόπο εμφάνισης των αποτελεσμάτων της χρησιμοποιεί έναν αλγόριθμο, με όνομα «Page Rank», ένα μηχανισμό που θα αναλυθεί εκτενέστερα στα επόμενα κεφάλαια της πτυχιακής. Το Page Rank είναι ένας αριθμός που προκύπτει από διάφορα μη κοινοποιημένα κριτήρια που έθεσαν οι προγραμματιστές της Google. Η πτυχιακή θα επικεντρωθεί κυρίως στο Page Rank καθώς και στην Google ως μηχανή αναζήτησης, για την ανάλυση των τεχνικών και των κανόνων αλλά και για τα παραδείγματα.

Μια αναζήτηση σε μια μηχανή αναζήτησης μπορεί να επιστρέψει χιλιάδες ως εκατομμύρια ιστοσελίδες. Ο αλγόριθμος της αναζήτησης βοηθά στην σωστή αξιολόγηση της κάθε σελίδας και την προβάλλει στην θέση που της αρμόζει. Πο-



ια είναι αυτά τα στοιχεία όμως που την επηρεάζουν; Τι μπορεί να κάνει μια ιστοσελίδα για να αξιολογηθεί καλύτερα και να εμφανιστεί υψηλότερα από τις άλλες;

Για την καλύτερη κατανόηση του προβλήματος αλλά και την εύρεση λύσης ακολουθεί ένα παράδειγμα. Έστω ότι ένας άνθρωπος στέκεται μέσα σε ένα πλήθος μερικών χιλιάδων ανθρώπων και κάποιος τον αναζητά, πώς θα τον βρει; Σε ένα πλήθος ανθρώπων όλοι φαίνονται ίδιοι. Έστω ότι υπάρχει ένα σύστημα όπου κατηγοριοποιεί το πλήθος των ανθρώπων. Για παράδειγμα, όλοι οι άντρες φοράνε μπλε ρούχα ενώ οι γυναίκες κόκκινα ρούχα. Αυτόματα μειώνει τους ανθρώπους που πρέπει να εξετάσει αυτός που αναζητά τον άνθρωπο στο μισό. Αν εφαρμοστούν και άλλα κριτήρια κατηγοριοποίησης το πλήθος γίνεται πιο ξεκάθαρο και σταδιακά μειώνεται η περιοχή αναζήτησης.

Στο παράδειγμα παρατηρείται ότι υπάρχει η ανάγκη ύπαρξης διαφόρων χαρακτηριστικών ώστε να γίνει η διάκριση μεταξύ όμοιων αντικειμένων. Χωρίς τα χαρακτηριστικά αυτά είναι πιο δύσκολο να εντοπιστεί ένας συγκεκριμένος άνθρωπος σε έναν όχλο, παρά απομονώνοντας τον σε υπό-ομάδες.

Η ιστοσελίδα είναι κάτι παραπάνω από έναν άνθρωπο που στέκεται σε ένα πλήθος. Χωρίς να ακολουθεί τους κανόνες του συστήματος είναι πολύ δύσκολο για κάποιον να την εντοπίσει γιατί πολύ απλά θα βρίσκεσαι μέσα στα χιλιάδες αποτελέσματα, όπως η 15<sup>η</sup> σελίδα αποτελεσμάτων, πράγμα μη θεμιτό. Έρευνες έχουν δείξει ότι οι χρήστες συνήθως διαβάζουν μέχρι και την 5<sup>η</sup> σελίδα αποτελεσμάτων όταν δεν βρίσκουν σχετικό με την αναζήτηση αποτέλεσμα.

Σύμφωνα με μια έρευνα που διεξήχθη το 2006 από την iProspect, μια από τις κορυφαίες εταιρίες που δραστηριοποιείται στον χώρο του μάρκετινγκ, και την Jupiter Research<sup>1</sup> έδειξε ότι το 62% των χρηστών των μηχανών αναζήτησης επισκέπτονται τις σελίδες που βρίσκονται στην πρώτη σελίδα των αποτελεσμάτων. Ένα άλλο πόρισμα της έρευνας είναι ότι το 90% βλέπει μόνο τα αποτελέσματα που βρίσκονται στις πρώτες τρεις σελίδες.

Επίσης, το 41% των χρηστών των μηχανών αναζήτησης όταν δεν βρίσκουν αυτό που ψάχνουν στην πρώτη σελίδα των αποτελεσμάτων αναζήτησης αναφέρουν ότι αλλάζουν τους όρους αναζήτησης ή ακόμη και μηχανή αναζήτη-

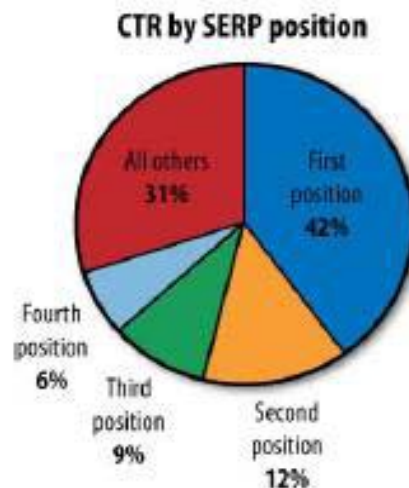
---

<sup>1</sup> iProspect, *Search Engine User Behavior Study*, April 2006

σης ενώ το 88% δηλώνει ότι θα προσπελάσει τις τρεις πρώτες σελίδες αποτελεσμάτων και μετά θα κάνει αλλαγές.

Τέλος, το 36% των χρηστών συμφωνούν ότι βλέποντας μια ιστοσελίδα στην κορυφή των αποτελεσμάτων, πιστεύουν ότι η σελίδα είναι η πιο κατάλληλη στον τομέα εφαρμογής της και δείχνουν μεγαλύτερη εμπιστοσύνη στο περιεχόμενό της, παρά σε άλλες σελίδες. Το φαινόμενο αυτό παρατηρείται κυρίως σε ιστοσελίδες εταιρειών.

Μια άλλη σημαντική έρευνα πάνω στο μάρκετινγκ, που αφορά της μηχανές αναζήτησης, που έγινε το 2006 πάνω στα αρχεία καταγραφής των αναζητήσεων της AOL δείχνει την σημαντικότητα της θέσης στα αποτελέσματα αναζήτησης στο ποσοστό των κλικ που λαμβάνει μια ιστοσελίδα. Ένα άλλο σημαντικό στοιχείο της έρευνας είναι ότι το 72% των χρηστών κάνουν κλικ στην πρώτη ιστοσελίδα των αποτελεσμάτων ενώ το 25,5% διαβάζουν όλη τη λίστα πρώτα και μετά αποφασίζουν ποια σελίδα θα επισκεφτούν. Στο διάγραμμα στην Εικόνα 1.4 φαίνεται καθαρά η διαφορά μεταξύ των κλικ του πρώτου αποτελέσματος από τις δεύτερες ή κατώτερες καταχωρήσεις.



Εικόνα 1.4: Στατιστικά των κλικ σε αποτελέσματα αναζήτησης

Σύμφωνα με τα παραπάνω, δημιουργείται η ανάγκη μια ιστοσελίδα να υπακούει σε διάφορους κανόνες, κανόνες που η κατάλληλη χρήση τους θα την κάνει να ξεχωρίσει. Η εύρεση των κανόνων και η βέλτιστη εφαρμογή τους σε όλη την έκταση της ιστοσελίδας ονομάζεται βελτιστοποίηση ιστοσελίδας για μηχανές αναζήτησης.

Το υπόλοιπο περιεχόμενο της πτυχιακής εργασίας διαμορφώνεται ως εξής:

Στο κεφάλαιο 2 παρουσιάζονται τα βασικά στοιχεία της βελτιστοποίησης ιστοσελίδων για μηχανές αναζήτησης. Γίνεται μια εισαγωγή στους όρους και στις εννοιές που συνδέονται άμεσα με την βελτιστοποίηση και μια αναφορά στους τομείς που αυτή εφαρμόζεται. Ακολουθεί η καταγραφή των στοιχείων και των τεχνικών της βελτιστοποίησης με παραδείγματα ενώ για πληρότητα καταγράφονται και τρόποι κατάχρησης της βελτιστοποίησης (κακόβουλη βελτιστοποίηση). Το κεφάλαιο 3 αναφέρει τις λύσεις βελτιστοποίησης που υπάρχουν ενώ το κεφάλαιο 4 παρουσιάζει διάφορα εργαλεία που χρησιμοποιούνται ευρέως για την βελτιστοποίηση μιας ιστοσελίδας. Στο κεφάλαιο 5 περιγράφεται η λειτουργικότητα καθώς και τεχνικές λεπτομέρειες της εφαρμογής seo που αναπτύχθηκε. Η πτυχιακή ολοκληρώνεται στο κεφάλαιο 6 όπου καταγράφονται τα γενικότερα συμπεράσματα και δίνονται κατευθυντήριες γραμμές για επεκτάσεις.



# 2 Βελτιστοποίηση Σελίδων για Μηχανές Αναζήτησης

Βελτιστοποίηση ιστοσελίδων ονομάζεται η διαδικασία διαμόρφωσης της δομής και του περιεχομένου μιας ιστοσελίδας σύμφωνα με ορισμένες τεχνικές και κανόνες ώστε να γίνει όσο το δυνατό πιο φιλική στις μηχανές αναζήτησης. Τελικός στόχος είναι η υψηλή κατάταξη της ιστοσελίδας στις σελίδες των αποτελεσμάτων των μηχανών αναζήτησης και η αύξηση της επισκεψιμότητας της μέσω οργανικών αποτελεσμάτων, δηλαδή αποτελεσμάτων χρηστών του διαδικτύου που ψάχνουν στις μηχανές αναζήτησης με τις λέξεις-κλειδιά που αφορούν το περιεχόμενο της ιστοσελίδας.

## 2.1 Βασικές Έννοιες & Ιστορική Αναδρομή

### 2.1.1 Λέξεις-Κλειδιά

Οι λέξεις-κλειδιά είναι ίσως το πιο σημαντικό στοιχείο που υπάρχει σε μια ιστοσελίδα. Οι λέξεις-κλειδιά δεν είναι τίποτα παραπάνω από κανονικές λέξεις ή φράσεις κειμένου. Ονομάζονται όμως κλειδιά γιατί αποτελούν το κλειδί μιας ιστοσελίδας. Είναι οι λέξεις που χαρακτηρίζουν το περιεχόμενο μιας ιστοσελίδας, λέξεις που με μια απλή καταγραφή θα πρέπει ο χρήστης να αντιλαμβάνεται περί τίνος πρόκειται η σελίδα. Οι μηχανές αναζήτησης είναι κύριοι χρήστες των λέξεων-κλειδιών και αποτελούν ζωτικό στοιχείο της βελτιστοποίησης.

Η αναζήτηση και η εύρεση των κατάλληλων λέξεων-κλειδιών είναι το πρώτο βήμα που πρέπει να ληφθεί σοβαρά υπόψη κατά τον σχεδιασμό μιας καμπάνιας βελτιστοποίησης. Η εύρεση των λέξεων-κλειδιών που θα χρησιμοποιηθούν θα πρέπει να απασχολήσει την ομάδα για περισσότερο από λίγες μέρες, αν θέλουν να πετύχουν το βέλτιστο αποτέλεσμα και την σωστή χρήση της δύναμης που κρύβουν. Χωρίς την σωστή μελέτη για το ποιες λέξεις είναι δημοφιλείς στην κατηγορία όπου ανήκει μια ιστοσελίδα, οι κόπτοι της καμπάνιας μπορεί να μην αποδώσουν ή θα αποδώσουν ελάχιστα. Η βελτιστοποίηση μιας

ιστοσελίδας χρειάζεται διάστημα μερικών εβδομάδων έως μήνες μέχρι να φέρει κάποια αποτελέσματα, οπότε πρέπει να κτιστεί πάνω σε μια σωστή βάση με τις κατάλληλες λέξεις-κλειδιά. Σε επόμενα κεφάλαια παρουσιάζονται και αναλύονται εργαλεία που βοηθούν στην εύρεση των πιο κατάλληλων λέξεων-κλειδιών καθώς και τον τρόπο παρακολούθησής τους.

Οι τρόποι εύρεσης των λέξεων-κλειδιών ποικίλουν. Ένας τρόπος είναι η ανάγνωση του περιεχόμενου μιας ιστοσελίδας και η εύρεση των σημαντικών λέξεων. Ύστερα, με την βοήθεια μιας μηχανής αναζήτησης, ελέγχονται ποιοτικά και ποσοτικά τα αποτελέσματα που προκύπτουν αναζητώντας κάθε σημαντική λέξη. Όσο περισσότερα είναι, τόσο σημαντική υποτίθεται είναι η λέξη. Ένας άλλος τρόπος είναι η αναζήτηση παρόμοιων ιστοσελίδων και υιοθέτηση μερικών λέξεων-κλειδιών που ήδη χρησιμοποιούνται. Τέλος, υπάρχουν εργαλεία που ο χρήστης εισάγει μια λέξη ή φράση και αυτά επιστρέφουν μια λίστα από προτεινόμενες λέξεις-κλειδιά καθώς και άλλα εργαλεία όπως το Google Adwords.

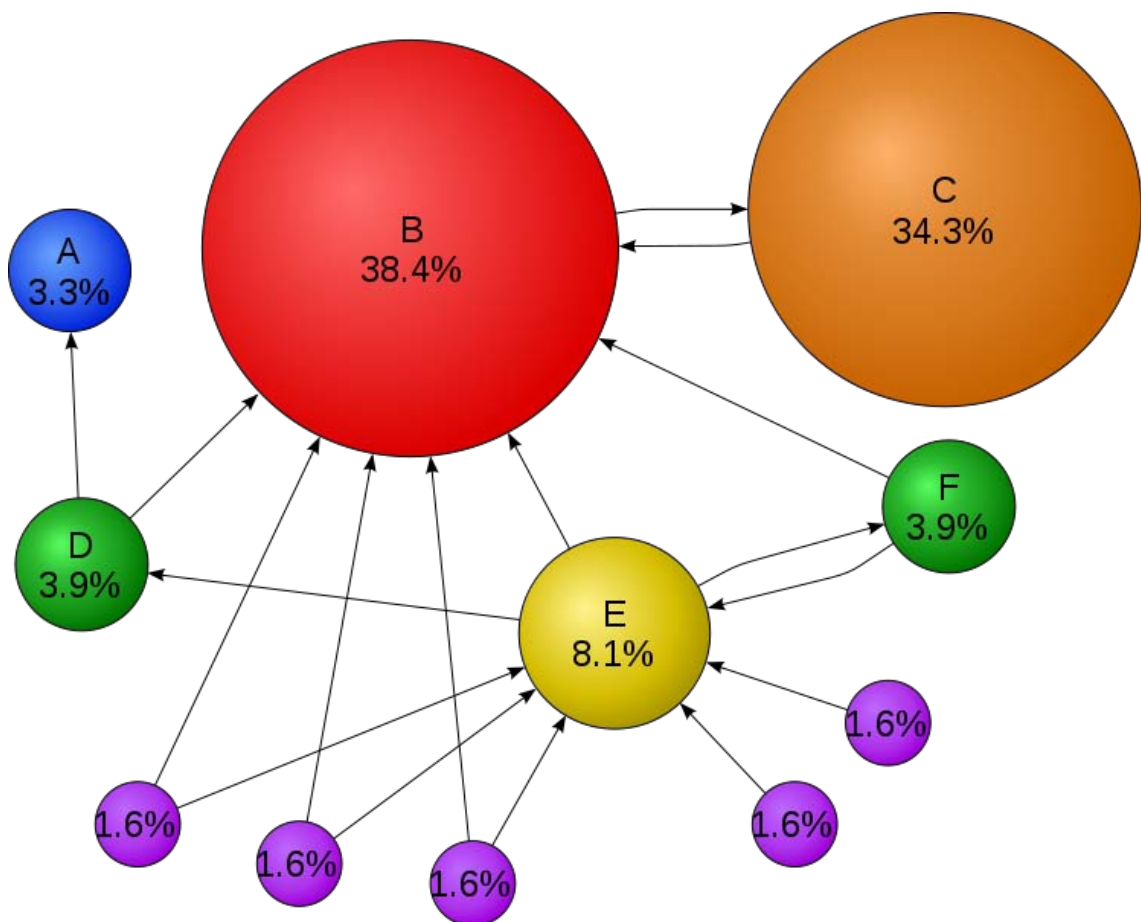
### **2.1.2 Page Rank**

Page Rank (PR) ονομάζεται ο αλγόριθμος ανάλυσης της ποιότητας των συνδέσμων μιας ιστοσελίδας. Ουσιαστικά είναι ένας αριθμός από το 0 έως το 10 που χαρακτηρίζει μια ιστοσελίδα και υπολογίζεται από μια εξίσωση με περισσότερα από 500 εκατομμύρια μεταβλητές και 2 εκατομμύρια όρους, σύμφωνα με την Google. Οι όροι είναι διάφορα στοιχεία που αφορούν την ποιότητα και σημαντικότητα των συνδέσμων. Όσο μεγαλύτερος είναι ο αριθμός PR, τόσο πιο ψηλά κατατάσσεται στα αποτελέσματα αναζήτησης η ιστοσελίδα και τόσο πιο αξιόπιστο και ποιοτικό είναι το περιεχόμενό της για το κοινό. Σύμφωνα με επίσημες πηγές της Google, το page rank κάθε ιστοσελίδας υπολογίζεται ανά τρεις μήνες περίπου, γεγονός που συνδράμει στην βραχυπρόθεσμη ορατότητα του αποτελέσματος μιας βελτιστοποίησης. Το όνομα «Page Rank» δόθηκε από τον δημιουργό του αλγορίθμου, Larry Page, και χρησιμοποιείται μόνο από την Google.

Σύμφωνα με την Google το PR στηρίζεται στη μοναδική δημοκρατική φύση του παγκόσμιου ιστού χρησιμοποιώντας την δομή των υπερσυνδέσεων ως μετρητή της αξίας μιας σελίδας. Η Google ερμηνεύει τον σύνδεσμο από την σελίδα A στην σελίδα B ως ψήφο της A προς την B.

Ένα παράδειγμα μέτρησης είναι το εξής. Έστω ότι υπάρχουν τρεις ιστοσελίδες A,B,Γ με PR 3,1 και 4 αντίστοιχα. Αν η ιστοσελίδα A έχει σύνδεση με την ιστοσελίδα B τότε η B λαμβάνει, σύμφωνα με την Google, ποιοτική ψήφο από την A. Αν η ιστοσελίδα Γ έχει και αυτή σύνδεση στην σελίδα B, τότε η ψήφος της Γ προς την B είναι πιο ποιοτική. Αυτό θα έχει ως αποτέλεσμα το PR της σελίδας B να ανέβει 1,2 μονάδες, ανάλογα τον άγνωστο, για το κοινό, αλγόριθμο. Οι σελίδες A και Γ θα μείνουν ανέπαφες.

Όπως απεικονίζεται στην Εικόνα 2.1, δεν μετριέται μόνο η ποσότητα των ψήφων προς μια ιστοσελίδα αλλά αναλύεται και η ποιότητα της ψηφοφόρου σελίδας. Η ποιοτική ψήφος μιας σημαντικής ιστοσελίδας, όπως η B στην C, έχει περισσότερο βάρος και προσδίδει μεγαλύτερη αξία σε άλλες σελίδες σε σχέση με τις ψήφους μικρότερης αξίας ιστοσελίδων, όπως οι μωβ στην E.



Εικόνα 2.1: Ποιοτικές και ποσοτικές συνδέσεις-ψηφοί ιστοσελίδων

Η πρωτοποριακή χρήση του PR από την Google έφερε επανάσταση στον τομέα της βελτιστοποίησης. Όλο και περισσότεροι αναλυτές άρχισαν να πειραματίζονται ώστε να ανακαλύψουν τι ακριβώς υπολογίζει η εξίσωση, ποια στοιχεία της σελίδας αφορά, τι θα πρέπει να βελτιστοποιήσουν σε μια σελίδα και με ποιον τρόπο μπορούν να επηρεάσουν την κατάταξη μιας ιστοσελίδας, με θεμιτές ή αθέμιτες μεθόδους. Τότε δημιουργήθηκαν εργαλεία για την μέτρηση και την πρόβλεψη του PR κάθε σελίδας καθώς και μισθωμένες υπηρεσίες βελτιστοποίησης.

Όσο η βελτιστοποίηση επηρεάζει την επιτυχία ιστοσελίδων που σημαίνει περισσότερα κέρδη και απόδοση, δεν ήταν δυνατό το Page Rank να παραμείνει άτρωτο. Διάφοροι ειδικοί του θέματος, γνωστοί και ως μαύρα πρόβατα της βελτιστοποίησης, προσπάθησαν να βρουν τρόπους για την χειραγώγηση του.

Ένας από αυτούς τους τρόπους ήταν η αγοραπωλησία συνδέσμων. Μια ιστοσελίδα με υψηλό Page Rank πουλούσε την εισαγωγή συνδέσμων από άλλες ιστοσελίδες μικρότερου PR στην δική της ιστοσελίδα με χρηματική αμοιβή. Με την χρήση αυτής της τεχνικής μια ιστοσελίδα χαμηλού PR μπορούσε να αγοράσει συνδέσμους σε υψηλού PR ιστοσελίδες και να επηρεάσει θετικά το δικό της PR ενώ η σελίδα-ευεργέτης αποκόμιζε οικονομικά κέρδη. Η διαδικασία αυτή όμως διέφθειρε την αξιοπιστία του PR για την αξιολόγηση μιας ιστοσελίδας γιατί ιστοσελίδες με μεγάλο κεφάλαιο μπορούσαν να εκτοξευτούν στις κορυφαίες θέσεις κατάταξης χωρίς να είναι απαραίτητα αξιόπιστες ή κατάλληλες. Το φαινόμενο αυτό δημιούργησε αθέμιτο χρηματικό ανταγωνισμό ανάμεσα σε ιστοσελίδες. Σήμερα η αγοραπωλησία έχει εξαλείφει και ο τρόπος υπολογισμού του page rank έχει αλλάξει.

Παρόμοιοι αλγόριθμοι κατάταξης είναι ο αλγόριθμος HITS, δημιούργημα του Jon Kleinberg για την μηχανή αναζήτησης Ask.com, ο IBM CLEVER project, και ο TrustRank (βλ. Παράρτημα).

### **2.1.3 Βελτιστοποίηση & Μάρκετινγκ (SEM)**

Μια ιστοσελίδα θεωρείται επιτυχής όταν έχει υψηλή επισκεψιμότητα. Τα υψηλά νούμερα επισκεπτών μεταφράζονται ως προσέλκυση περισσότερων διαφημιστών και πελατών, ανάλογα το αντικείμενο που διαπραγματεύεται μια ιστοσελίδα.



Αντικείμενο εμπορικής αξίας μπορεί να είναι ένα προϊόν ή μια υπηρεσία που παρέχει η ιστοσελίδα αλλά και ο χώρος διαφήμισης. Πολλές ιστοσελίδες λειτουργούν μη κερδοσκοπικά, παρέχουν περιεχόμενο όπως άρθρα ενημέρωσης ή προσφορά γνώσης χωρίς αντίτιμο. Στόχος αυτών των ιστοσελίδων είναι η προσέλκυση διαφημιστών ώστε να αποκτήσουν έσοδα για την κάλυψη των αναγκών τους.

Οι επισκέπτες μιας ιστοσελίδας χωρίζονται σε άμεσους, επισκέπτες που την γνωρίζουν και την επισκέπτονται απ' ευθείας με την χρήση της διεύθυνσης της, και στους επισκέπτες που την βρίσκουν ως σύνδεσμο σε μια άλλη ιστοσελίδα. Οι σύνδεσμοι αυτοί μπορεί να είναι διαφημίσεις ή εθελοντικές συνδέσεις, παραπομπές ή αυτοματοποιημένες διαδικασίες. Οι χρήστες μιας μηχανής αναζήτησης ανήκουν στην δεύτερη κατηγορία. Ένα μεγάλο μέρος των επισκεπτών μιας ιστοσελίδας προέρχεται από μηχανές αναζήτησης, γεγονός που οδηγεί στο συμπέρασμάς ότι η επιτυχία μιας ιστοσελίδας βασίζεται σε μεγάλο βαθμό από την κατάταξη της στις μηχανές αναζήτησης.

Ο όρος SEM είναι ακρωνύμιο των λέξεων *Search Engine Marketing* και αναφέρεται σε μια μορφή του Διαδικτυακού Μάρκετινγκ όπου αντικείμενο της είναι η αναζήτηση τρόπων προώθησης των ιστοσελίδων ώστε να βελτιώσουν την κατάταξη και την ορατότητα τους στις σελίδες των αποτελεσμάτων των μηχανών αναζήτησης. Μερικοί από τους τρόπους αυτούς είναι η προσθήκη συνδέσμων επί πληρωμής και η στοχευμένη διαφήμιση.

Με τον παραπάνω τρόπο άρχισε να συμβάλει η βελτιστοποίηση των ιστοσελίδων στο Μάρκετινγκ. Πολλές ιστοσελίδες στηρίζουν τα έσοδα τους στην επιτυχία τους, γι αυτό και επενδύουν αξιόλογα ποσά στην βελτιστοποίηση τους. Η καλύτερη βελτιστοποίηση φέρνει περισσότερα έσοδα. Ως αποτέλεσμα, δημιουργήθηκαν επαγγέλματα και θέσεις εργασίας σε εταιρίες κατασκευής ιστοσελίδων όπως βελτιστοποιητές ιστοσελίδων αλλά και εταιρίες που παρέχουν αποκλειστικά μισθωμένες υπηρεσίες βελτιστοποίησης. Η επιτυχία στον τομέα της βελτιστοποίησης αυτόματα σημαίνει και οικονομική επιτυχία. Εταιρίες που ασχολούνται με την βελτιστοποίηση και έχουν επιτυχημένες ιστοσελίδες στο δυναμικό τους είναι περιζήτητες και ακριβοπληρωμένες.

## 2.1.4 PPC ή Βελτιστοποίηση

Ο αυξανόμενος αριθμός των καταναλωτών που αγοράζουν προϊόντα και υπηρεσίες online έχει οδηγήσει πολλές εταιρίες να στραφούν και να χρησιμοποιήσουν το SEM και να προσπαθούν να προσελκύσουν περισσότερους χρήστες στην ιστοσελίδα τους ώστε να έχουν το προνόμιο στην αυξανόμενη αγορά.

Υπάρχουν δύο κύριες μορφές του SEM που μπορούν να χρησιμοποιηθούν, η διαφήμιση *Pay Per Click* και η βελτιστοποίηση της ιστοσελίδας. Οι δυο αυτοί όροι είναι εντελώς διαφορετικοί. Πολλοί πιστεύουν ότι βελτιστοποίηση ιστοσελίδας για μηχανές αναζήτησης σημαίνει αγορά PPC καμπάνιας διότι αρκετές εταιρίες υποστηρίζουν ότι προσφέρουν υπηρεσίες βελτιστοποίησης ενώ στην ουσία προσφέρουν υπηρεσίες PPC. Ο λόγος που το συμβαίνει αυτό είναι η άγνοια όσο αφορά τον πελάτη και η PPC είναι άκοπη και πιο κερδοφόρα για την εταιρία.

Το PPC είναι μια μορφή διαφήμισης της ιστοσελίδας σε μια μηχανή αναζήτησης. Ο χρήστης αναζητά κάποιο όρο στην μηχανή αναζήτησης και τα αποτελέσματα, που η δεύτερη επιστρέφει, περιέχουν σε εμφανές σημείο διαφημίσεις κειμένου με ιστοσελίδες σχετικά με τον όρο. Ο χρήστης επισκέπτεται την ιστοσελίδα από τον σύνδεσμο της στα αποτελέσματα και ταυτόχρονα η μηχανή αναζήτησης χρεώνει τον ιστότοπο ένα πόσο που αντιστοιχεί στην συγκεκριμένη λέξη-κλειδί.

Ένας ιδιοκτήτης ιστοσελίδας, που επιθυμεί να προωθήσει την ιστοσελίδα του με PPC, επισκέπτεται τις μηχανές αναζήτησης που επιθυμεί να διαφημιστεί και εγγράφεται στην ειδική υπηρεσία που παρέχει η έκαστη μηχανή αναζήτησης. Στην φόρμα εγγραφής ο χρήστης καλείται να επιλέξει σε ποιες λέξεις-κλειδιά θα εμφανίζεται η ιστοσελίδα του.

Όπως είναι λογικό, οι λέξεις-κλειδιά δεν αξίζουν το ίδιο. Μια λέξη-κλειδί που δεν χρησιμοποιείται συχνά θα είναι πιο φθηνή από μια λέξη-κλειδί που χρησιμοποιείται καθημερινά. Για παράδειγμα, η λέξη «business» μπορεί να χρεώνεται μέχρι και 2€ ανά κλικ.

Αφού επιλέξει τις επιθυμητές λέξεις-κλειδιά, η μηχανή αναζήτησης πιστώνει τον διαφημιζόμενο ένα ποσό από το οποίο θα αφαιρείται η αξία της λέξης-κλειδί από την οποία δέχτηκε επισκέπτη η ιστοσελίδα του. Όταν το ποσό του

διαφημιζόμενου μηδενίσει, η καμπάνια του απενεργοποιείται και ο χρήστης ενημερώνεται για την πίστωση νέου ποσού ή διαγραφή του λογαριασμού του.

Η πιο γνωστή μηχανή αναζήτησης που λειτουργούσε μόνο με PPC ήταν η Overture, η οποία εξαγοράστηκε από την Yahoo! για εμπορική εκμετάλλευση και ενσωμάτωση της στην Yahoo! Search. Η πιο γνωστή υπηρεσία αυτή την περίοδο είναι το Google AdWords από την μηχανή αναζήτησης Google.

Η χρήση του PPC παρέχει ευκολίες προώθησης σε μια ιστοσελίδα. Καταρχάς, δεν χρειάζονται ειδικές γνώσεις προγραμματισμού ή βελτιστοποίησης γιατί η ιστοσελίδα δεν θα αλλάξει ούτε δομή, ούτε περιεχόμενο. Χρειάζεται μόνο ένας γνώστης του αντικειμένου του SEM να ετοιμάσει την καμπάνια με τις επιθυμητές λέξεις-κλειδιά. Δεύτερον, τα αποτελέσματα της καμπάνιας θα φανούν πολύ πιο γρήγορα από μια καμπάνια βελτιστοποίησης. Τρίτον, η εταιρία χρεώνεται μόνο για τους επισκέπτες που θα πατήσουν την διαφήμιση της και μπορεί να επιλέξει σε ποιες λέξεις-κλειδιά επιθυμεί να εμφανίζεται. Τέλος, δίνονται πολλές δυνατότητες για την επιλογή του κοινού εμφάνισης της διαφήμισης, όπως το ποιες σελίδες θα εμφανίζονται καθώς και ποιες γεωγραφικές περιοχές στοχεύονται στην διαφήμιση.

Από την άλλη πλευρά η χρήση της PPC προώθησης έχει ημερομηνία λήξης. Σε αντίθεση με την PPC, η βελτιστοποίηση μιας ιστοσελίδας έχει πιο μακρόχρονα αποτελέσματα ενώ της PPC εξασθενούν. Το PPC επίσης κοστίζει αρκετά σε μια εταιρία αν θέλει να ανταγωνιστεί τις υπόλοιπες του χώρου. Όπως σε άλλους τομείς του εμπορίου και του ανταγωνισμού, έτσι και στον τομέα προώθησης, το κεφάλαιο μιας ιστοσελίδας παίζει σημαντικό ρόλο στην προώθηση της. Όσο μεγαλύτερο ποσό διατεθεί για διαφήμιση, τόσο αυξάνονται οι πιθανότητες εμφάνισης της διαφήμισης και η επιτυχία της. Τέλος μια καμπάνια PPC αναφέρεται σε μια μόνο μηχανή αναζήτησης. Για περισσότερες μηχανές το κόστος είναι απαγορευτικό αν το κεφάλαιο είναι σχετικά μικρό.

Τέλος, δεν πρέπει να μην υπολογίζεται ο παράγοντας χρήστης. Η διαφήμιση από πολλούς θεωρείται ότι είναι ένας τρόπος προώθησης ενός προϊόντος που ενδέχεται να προσφέρει λιγότερα από ότι υπόσχεται. Το ίδιο ισχύει και στις διαφημίσεις που εμφανίζονται στα αποτελέσματα των μηχανών αναζήτησης. Οι χρήστες εμπιστεύονται περισσότερο τις τρεις πρώτες ιστοσελίδες που εμφανί-

ζονται στα αποτελέσματα μιας αναζήτησης διότι αυτές έχουν περάσει από έλεγχο και είναι πλέον αποδεκτές ως οι καταλληλότερες για την λέξη-κλειδί που αναζητούν, παρά τις διαφημίσεις που κρύβουν από πίσω οικονομικά οφέλη.

Παρόλο που το PPC & η βελτιστοποίηση είναι διαφορετικά αντικείμενα και ανταγωνίζονται μεταξύ τους, μπορούν να συνεργαστούν άψογα. Για παράδειγμα, μια εταιρεία λανσάρει ένα προϊόν στο διαδίκτυο. Αφού στήσει το ηλεκτρονικό της κατάστημα, ξεκινά την διαδικασία βελτιστοποίησης. Παράλληλα, ξεκινάει και μια καμπάνια PPC σε μια ή δυο μηχανές αναζήτησης. Επενδύει στην καμπάνια περίπου τρεις μήνες ενώ παράλληλα προσπαθεί να εμφανιστεί και ως αποτέλεσμα αναζήτησης, άσχετο από την διαφημιστική περιοχή. Αφού το πετύχει, απενεργοποιεί την καμπάνια. Με τον τρόπο αυτό καλύπτεται ο χρόνος που απαιτείται για την αξιολόγηση μιας σελίδας από μια μηχανή αναζήτησης και εμφάνιση της σε αξιολογη θέση.

### **2.1.5 Καλή και Κακόβουλη Βελτιστοποίηση**

Όπως αναφέρθηκε παραπάνω, η βελτιστοποίηση είναι ένα από τα επικρατέστερα σημεία ανταγωνισμού μεταξύ των ιστοσελίδων σε θέματα μάρκετινγκ και όχι μόνο. Ως ακόλουθο, υπάρχει η καλή και η κακόβουλη βελτιστοποίηση και δυο στρατόπεδα, των μηχανών αναζήτησης και των κακόβουλων βελτιστοποιητών.

Η καλή βελτιστοποίηση, ή αλλιώς white hat SEO, είναι θεμιτή από τις μηχανές αναζήτησης και προσφέρει μακρόχρονα επιτυχία και σταθερότητα στα αποτελέσματα αναζήτησης. Αποτελείται από μια σειρά κανόνων που χρησιμοποιούν οι μηχανές αναζήτησης για την αξιολόγηση του περιεχόμενου των ιστοσελίδων. Οι κανόνες τροποποιούνται ή καταργούνται με την πάροδο του χρόνου. Κύριος λόγος της συνεχούς αλλαγής είναι η μη επάρκεια τους ή η κατάχρησή τους από το στρατόπεδων των κακόβουλων βελτιστοποιητών.

Η κακόβουλη βελτιστοποίηση, ή αλλιώς black hat SEO, δεν διαφέρει σημαντικά από τις καλές πρακτικές που αναφέρονται παραπάνω. Η αντίθεση τους βρίσκεται στην κατάχρηση των κανόνων και των τεχνικών. Η κατάχρηση των κανόνων φέρει πολύ καλά αποτελέσματα μικρής διάρκειας και ασταθή. Αυτό οφείλεται στο γεγονός ότι οι μηχανές αναζήτησης αμύνονται στις καταχρήσεις. Η

χρήση κακών τεχνικών δεν συνιστάται και διώκεται από τις μηχανές αναζήτησης. Μια μηχανή αναζήτησης διατηρεί μαύρη λίστα και προσθέτει σε αυτήν όσες σελίδες ανακαλυφθούν ότι καταχράζονται τις τεχνικές βελτιστοποίησης. Ακολουθεί ένα παράδειγμα κακόβουλης βελτιστοποίησης από το παρελθόν, προτού οι μηχανές αναζήτησης απορρίψουν τα μεταδεδομένα.

Έστω ένας webmaster θέλει να προωθήσει μια ιστοσελίδα με πολιτική προπαγάνδα. Με την χρήση των meta tags θα μπορούσε να δηλώσει το περιεχόμενο της ιστοσελίδας του ως μια σελίδα που αφορά τον Michael Jackson. Αυτό επιτυγχάνεται με την χρήση των κατάλληλων λέξεων-κλειδιών και κρυφό κείμενο πάνω στην σελίδα για την επανάληψη τους. Ένας αλγόριθμος αναζήτησης που χρησιμοποιεί μόνο τα μεταδεδομένα θα έδινε στην σελίδα υψηλή κατάταξη στα αποτελέσματα αναζήτησης του όρου «Michael Jackson» διότι θα στηριζόταν στο γεγονός ότι οι λέξεις Michael Jackson βρίσκονται πολλές φορές στο κείμενο της σελίδας. Όταν ο χρήστης αναζητούσε το λήμμα «Michael Jackson» θα έβγαινε πρώτη η σελίδα της πολιτικής προπαγάνδας, με αποτέλεσμα ο χρήστης να μην είναι ευχαριστημένος από το αποτέλεσμα διότι δέχεται ιστοσελίδες άσχετες με το θέμα που αναζητά. Αυτό θα συνεχιζόταν μέχρι να ανακαλύψουν οι προγραμματιστές της μηχανής αναζήτησης ότι γίνεται κατάχρηση του παράγοντα meta tag και να δημιουργήσουν κάποιο μπάλωμα στον αλγόριθμο αναζήτησης ή να βάλουν την σελίδα στην μαύρη λίστα και να την αποκλείσουν από τα ευρετήρια της μηχανής αναζήτησης.

### **2.1.6 Φιλικές διευθύνσεις**

Όπως το περιεχόμενο μιας ιστοσελίδας θα πρέπει να είναι ευανάγνωστο από μια μηχανή αναζήτησης για καλύτερη αξιολόγηση, έτσι και η διεύθυνση της σελίδας θα πρέπει να την χαρακτηρίζει. Οι φιλικές διευθύνσεις, γνωστές και ως SEF από τα ακρωνύμια των λέξεων *Search Engine Friendly URL* συνδέονται με την βελτιστοποίηση ιστοσελίδων. Οι φιλικές διευθύνσεις αναφέρονται σε μια μορφή διεύθυνσης, πιο ευανάγνωστη για τον χρήστη αλλά και την μηχανή αναζήτησης καθώς και πιο περιγραφική.

Μια φιλική διεύθυνση είναι μια διεύθυνση σχετική με το περιεχόμενο της σελίδας που αντιπροσωπεύει. Για παράδειγμα, αν υπάρχει μια σελίδα που περιγράφει το φρούτο λεμόνι, η διεύθυνση */frouta/lemoni.html* είναι πιο φιλική σε

μια μηχανή αναζήτησης και αντικατοπτρίζει το περιεχόμενο της ιστοσελίδας παρά μια διεύθυνση του τύπου */a/b.html*. Στις στατικές σελίδες η προσαρμογή των διευθύνσεων είναι μια εύκολη διαδικασία που αρκείται στην δημιουργία φακέλων και στατικών html σελίδων ανάλογες με το περιεχόμενό τους.

Το πρόβλημα με τις δυσανάγνωστες διευθύνσεις συναντάται κυρίως στις δυναμικές ιστοσελίδες. Μια δυναμική ιστοσελίδα δεν έχει κάποια πραγματική διεύθυνση που απεικονίζει ένα αρχείο αλλά αποτελείται από δυναμικές διευθύνσεις. Μια δυναμική διεύθυνση συνθέτεται από το αρχείο, όπου εκτελείται ένα κομμάτι κώδικα, και διάφορες παραμέτρους οι οποίες χρησιμοποιούνται για την κλήση του ανάλογου περιεχομένου. Μια δυναμική διεύθυνση θα μπορούσε να είναι της μορφής *frouta.php?id=5* ή *frouta.php?name=lemoni*. Όπως ένας απλός χρήστης, δεν αντιλαμβάνεται εύκολα σε τι αναφέρεται μια σελίδα με διεύθυνση *id=5*, έτσι και μια μηχανή αναζήτησης δεν μπορεί να καταλάβει τι περιεχόμενο θα συναντήσει μέσα στην σελίδα. Η μεθοδολογία μετατροπής των διευθύνσεων βασίζεται σε έναν μηχανισμό μετασχηματισμού διευθύνσεων, ο οποίος μετασχηματίζει το *frouta.php?name=lemoni* ή *id=5* σε */frouta/5* ή ακόμη καλύτερα */frouta/lemoni/* και αναλύεται στα επόμενα κεφάλαια.

### **2.1.7 Βελτιστοποίηση Σελίδων για Κινητά Τηλέφωνα**

Ο διαδικτυακός κόσμος μεταφέρεται σταδιακά από της ηλεκτρονικούς υπολογιστές στα κινητά τηλέφωνα και το διαδίκτυο μέσω κινητού γίνεται η νέα μόδα της εποχής. Οι φυλλομετρητές και τα έξυπνα τηλέφωνα μετέτρεψαν τον παραδοσιακό κινητό κόσμο σε ολόκληρο σύμπαν από ιστοσελίδες, καταστήματα, υπηρεσίες και όχι μόνο. Αυτό σημαίνει ότι αν μια ιστοσελίδα θέλει να εκμεταλλευτεί στο έπακρον της πηγές επισκεπτών θα πρέπει να ακολουθήσει την νέα τάση. Όπως όλα μεταφέρονται στον μικρόκοσμο των κινητών, το ίδιο συμβαίνει και με την βελτιστοποίηση ιστοσελίδων, μεταφέρεται και αποκτά νέα μορφή.

Οι πιο διάσημες ιστοσελίδες, όπως το Google, το Yahoo!, το Facebook, το Twitter και πολλά ακόμη έχουν μεταφερθεί ήδη στον μικρόκοσμο των κινητών τηλεφώνων και πολλές υπηρεσίες είναι σε δοκιμαστικό στάδιο όπως m-Government, m-Banking και όχι μόνο.

Η διαδικασία αναζήτησης πληροφοριών είναι παρόμοια με αυτή των ηλεκτρονικών υπολογιστών. Οι χρήστες επισκέπτονται τις μηχανές αναζήτησης για να αναζητήσουν ορισμένες πληροφορίες. Όταν μια ιστοσελίδα είναι συμβατικά βελτιστοποιημένη δεν σημαίνει αυτόματα ότι είναι βέλτιστη και για κινητά. Πολλοί από τους κανόνες βεβαίως παραμένουν ίδιοι, κόλπα κατάχρησης επίσης, αλλά συναντώνται ορισμένες διαφοροποιήσεις.

Καταρχάς μια ιστοσελίδα θα πρέπει να είναι ανοιχτή στις μηχανές αναζήτησης κινητών τηλεφώνων. Όταν δημιουργείται μια ιστοσελίδα πρέπει να δοθούν ορισμένα μεταδεδομένα έτσι ώστε να μπορεί να αναγνωσθεί. Αντίστοιχα, στις ιστοσελίδες των κινητών τηλεφώνων θα πρέπει να γίνουν παρόμοιες προσθήκες μεταδεδομέδων με μικρές διαφοροποιήσεις στην διαδικασία.

Η διαδικασία είναι απλή και λιγότερο απαιτητική από την συμβατική βελτιστοποίηση. Η ιστοσελίδα θα πρέπει να καταχωρηθεί χειροκίνητα στις μηχανές αναζήτησης. Όσες μηχανές αναζήτησης διαθέτουν αντίστοιχο ευρετήριο για ιστοσελίδες κινητών τηλεφώνων, έχουν ειδική επιλογή για την καταχώρηση ιστοσελίδας συμβατής με αυτό. Με την πραγματοποίηση της καταχώρησης, έχει επιτευχθεί το 50% της βελτιστοποίησης της ιστοσελίδας για κινητά τηλέφωνα.

Ένα καινούριο στοιχείο που συναντάται στην βελτιστοποίηση ιστοσελίδων για κινητά τηλέφωνα είναι το `metatxt`. Η ιδέα του `metatxt` βασίζεται στην εύκολη ανάγνωση ενός αρχείου κειμένου απευθείας από τους spiders χωρίς να χρειάζεται να εξετάσουν όλη την σελίδα. Το `metatxt` περιέχει πληροφορίες σχετικά με την ιστοσελίδα, πληροφορίες που είναι χρήσιμες όχι μόνο στις μηχανές αναζήτησης αλλά και σε άλλες υπηρεσίες ή επεκτάσεις των ιστοσελίδων, ακόμη και στους χρήστες.

Προς το παρόν ο τομέας της βελτιστοποίησης για κινητά τηλέφωνα είναι νέος και το `metatxt` αποτελεί μια πρώτη κίνηση. Όμως, υπάρχουν ενδοιασμοί και σκέψεις ότι θα υπάρξει φαινόμενο αντίστοιχο με τα `meta tags`, όσο αφορά την ασάφεια τους και την εκμετάλλευσή τους από επιτήδειους. Όπως κάθε νέος τομέας συναντά δυσκολίες και προβλήματα στα πρώτα στάδια του και αργότερα επιλύονται, αντίστοιχα στην βελτιστοποίηση ιστοσελίδων για κινητά τα προβλήματα ενδέχεται να αντιμετωπιστούν σύντομα.

### 2.1.8 Βελτιστοποίηση Σελίδων και Ελληνική Γλώσσα

Εδώ και πολλά χρόνια, η Αγγλική γλώσσα είναι κυρίαρχη σε θέματα τεχνολογίας, διαδικτύου και όχι μόνο. Τα αγγλικά συναντώνται καθημερινά από τις απλές ηλεκτρικές συσκευές μέχρι και στον τύπο. Μπορεί στην καθημερινότητα να μην προκαλούν μεγάλη σύγχυση, στην βελτιστοποίηση ιστοσελίδων και τις μηχανές αναζήτησης όμως δημιουργούν μεγάλα προβλήματα που θα αργήσει να δοθεί λύση.

Ως γνωστόν, στις ιστοσελίδες χρησιμοποιείται ο γραπτός λόγος. Ο λόγος χρησιμοποιείται από τις μηχανές αναζήτησης ώστε να αντλήσουν τα δεδομένα από μια ιστοσελίδα και να την κατατάξουν αντίστοιχα. Οι αγγλικοί, λατινικοί, χαρακτήρες είναι ενσωματωμένοι σε κάθε υπολογιστή και γίνονται αντιληπτά από τα περισσότερα προγράμματα. Τι γίνεται όμως με τις πολυτονικές γλώσσες ή γλώσσες με διαφορετικό, του λατινικού, αλφάβητο; Παρακάτω θα αναλυθούν δυο περιπτώσεις προβλημάτων που αντιμετωπίζει η ελληνική γλώσσα.

Το πρόβλημα της γλώσσας και η απόδοση του περιεχομένου λύθηκε, ευτυχώς, σχετικά νωρίς. Οι μεγάλες μηχανές αναζήτησης, στην αρχή, δεν αναγνώριζαν ιστοσελίδες εκτός των αγγλικών, λόγω της αδυναμίας των spiders να αντιληφθούν διαφορετικές γλώσσες. Ως αποτέλεσμα, οι ελληνικές ιστοσελίδες εντάσσονταν μόνο σε ελληνικές μηχανές αναζήτησης, οι οποίες δεν ήταν κάτι παραπάνω από ένας κατάλογος όπου ο χρήστης πρόσθετε χειροκίνητα την ιστοσελίδα του, και ορισμένες πληροφορίες για αυτήν, και ο κατάλογος είχε δυνατότητα αναζήτησης στα περιεχόμενα του. Ιστοσελίδες που δεν ήταν καταχωρημένες στους καταλόγους, ήταν χαμένες στο πέλαγος του διαδικτύου. Αργότερα, άρχισαν να κατασκευάζουν πολυγλωσσικούς spiders για την αναγνώριση μη λατινογενών γλωσσών.

Ένα άλλο πρόβλημα που υφίσταται μέχρι σήμερα είναι ο γραπτός λόγος που ποικίλει από γλώσσα σε γλώσσα. Η ελληνική γλώσσα έχει την ιδιαιτερότητα της κλίσης των ρημάτων, των ουσιαστικών και των επιθέτων, ιδιαιτερότητα που δυσχεράνει προγραμματιστές και μη. Στην αγγλική γλώσσα ένα αντικείμενο χαρακτηρίζεται από μια λέξη. Για παράδειγμα, το παιδί λέγεται «kid» Όπως και να χρησιμοποιηθεί η λέξη θα παραμείνει αμετάβλητη. Στα ελληνικά μπορεί να δοθεί ως «παιδί», «παιδιού» «παιδιά» Για έναν υπολογιστή οι τρεις προηγούμε-



νες λέξεις αντιπροσωπεύουν τρία διαφορετικά πράγματα, γεγονός που δεν αληθεύει. Το φαινόμενο αυτό φέρνει δυσκολία στην κατανόηση αλλά και την σωστή χρήση των λέξεων-κλειδιών. Οι Η/Υ δεν έχουν ανεπτυγμένη τεχνίτη νοημοσύνη ώστε να αντιλαμβάνονται τις λέξεις ως μια και όχι διαφορετικές. Γίνονται προσπάθειες στον τομέα αυτό, αλλά δεν υπάρχει κάποια σημαντική εξέλιξη γιατί ο όγκος των λέξεων αλλά και ο αριθμός των γλωσσών είναι τεράστιος. Αν και η ονομαστική επικρατεί, πολλοί βελτιστοποιητές ιστοσελίδων λαμβάνουν υπόψη και άλλες κλίσεις ουσιαστικών.

Ένα πιο πρόσφατο θέμα αφορά την χρήση ελληνικών χαρακτήρων σε ονόματα αρχείων ή φιλικών διευθύνσεων στο διαδίκτυο. Παλαιότερα δεν υπήρχε λόγος αναφοράς στο θέμα γιατί τα ελληνικά ως όροι σε προγράμματα ήταν σχεδόν απαγορευμένο.

Όπως αναφέρεται και στο κεφάλαιο με τους κανόνες και τις τεχνικές της βελτιστοποίησης, μια διεύθυνση με αναφορά κάποιας λέξης-κλειδί, ακόμη και στην διαδρομή ενός αρχείου, μετράει ως θετικός παράγοντας για την σελίδα. Όμως, οι ελληνικές ιστοσελίδες χρησιμοποιούν ελληνικές λέξεις-κλειδιά. Αν το όνομα στην διεύθυνση είναι στην αγγλική γλώσσα, χάνεται ένας σημαντικός παράγοντας. Τι λύση μπορεί να δοθεί;

Παρόλο που οι μεγαλύτερες μηχανές αναζήτησης πλέον αναγνωρίζουν τους ελληνικούς χαρακτήρες στις διευθύνσεις συνδέσμων και αντιλαμβάνονται τις λέξεις-κλειδιά, δεν υπάρχει πλήρη υποστήριξη και αναγνώριση των ελληνικών χαρακτήρων από όλους τους φυλλομετρητές και τους ηλεκτρονικούς υπολογιστές, με αποτέλεσμα η διεύθυνση να μετατρέπεται σε κάτι εντελώς διαφορετικό και μη φιλικό για τον χρήστη.

Ως παράδειγμα θα χρησιμοποιηθεί η φιλική διεύθυνση ενός άρθρου σχετικά με την ελληνική γλώσσα και την utf-8 κωδικοποίηση. Η διεύθυνση του άρθρου είναι η εξής.

*<http://www.webdevzone.gr/arthra/ελληνικά-και-utf-8>*

Στην διεύθυνση διακρίνεται καθαρά ότι είναι φιλική διεύθυνση αφού δεν περιέχει ειδικούς χαρακτήρες και ο χρήστης αντιλαμβάνεται ότι θα οδηγηθεί σε μια σελίδα όπου έχει ένα άρθρο σχετικά με την ελληνική γλώσσα και την utf-8 κωδικοποίηση.

Το πρωτόκολλο URL όμως χρησιμοποιεί ISO latin κωδικοποίηση και δεν μπορεί να διαβάσει μη λατινικούς χαρακτήρες, οπότε μετατρέπει την διεύθυνση σε δεκαεξαδική μορφή που βλέπετε παρακάτω χωρίς να καταστραφεί η σύνδεση.

*<http://www.webdevzone.gr/arthra/%CE%B5%CE%BB%CE%BB%CE%B7%CE%BD%CE%B9%CE%BA%CE%AC-%CE%BA%CE%B1%CE%B9-utf8>*

Παρατηρείται ότι οι χαρακτήρες τύπου %CE%B5 και λοιποί δεν βγάζουν κάποιο νόημα. Αν ακολουθηθεί ο σύνδεσμος ο χρήστης θα μεταφερθεί στην ίδια ακριβώς σελίδα με αυτήν που θα έγγραφε ο ίδιος την διεύθυνση. Δεν χάνεται η ακεραιότητα της υπερσύνδεσης αλλά δεν βρίσκονται κάποια κλειδιά στην διεύθυνση που θα βοηθήσουν στην βελτιστοποίηση της ιστοσελίδας.

Η διόρθωση του φαινομένου αυτό οφείλεται στην αναγνώριση ή μετατροπή των λατινικών χαρακτήρων σε μη. Για παράδειγμα, πριν μερικούς μήνες, δεν υποστήριζαν όλοι οι φυλλομετρητές utf-8 κωδικοποίηση στο URL αλλά iso latin, όπως ακριβώς και το ίδιο το πρωτόκολλο. Οι πιο πρόσφατες εκδόσεις τους αναγνωρίζουν την utf κωδικοποίηση και απλά την μετασχηματίζουν σε iso latin για να γίνει σωστά η υπερσύνδεση. Παρόμοιο είναι το φαινόμενο με τις μεθόδους των γλωσσών προγραμματισμού.

Το παραπάνω φαινόμενο έχει δημιουργήσει δυο στρατόπεδα βελτιστοποιητών σελίδων, που το καθένα υποστηρίζει την δική του άποψη στο αν θα πρέπει ή όχι να χρησιμοποιούνται ελληνικοί χαρακτήρες στις διευθύνσεις. Οι πρώτοι χρησιμοποιούν τα ελληνικά όπως ακριβώς θα χρησιμοποιούσαν τα αγγλικά και υποστηρίζουν ότι αφού οι τοπικές μηχανές αναζήτησης αναγνωρίζουν τις λέξεις-κλειδιά στις διευθύνσεις θα πρέπει να χρησιμοποιούνται, ενώ οι δεύτεροι υποστηρίζουν ότι πρέπει να λαμβάνεται υπόψη και ο παράγοντας χρήστης, όχι μόνο η βελτιστοποίηση, και να χρησιμοποιούν διευθύνσεις που δεν θα προκαλέσουν κάποιο πρόβλημα όπως η γραφή σε greeklish.

### **2.1.9 Ιστορική Αναδρομή**

Οι ιδιοκτήτες ιστοσελίδων ξεκίνησαν να βελτιστοποιούν τις ιστοσελίδες για τις μηχανές αναζήτησης από τα μέσα της δεκαετίας του '90, μιας και οι πρώτες μηχανές αναζήτησης καταχωρούσαν ήδη τον Παγκόσμιο Ιστό στα ευρετήρια τους.

Αρχικά ένας webmaster όφειλε να προσθέσει την διεύθυνση της ιστοσελίδας του σε διάφορες μηχανές αναζήτησης αν ήθελε να εμφανίζεται στα αποτελέσματά τους. Οι μηχανές αναζήτησης με την σειρά τους έστελναν έναν spider για να ανιχνεύσει σελίδα, να εξάγει τις υπερσυνδέσεις προς άλλες σελίδες και να επιστρέψει τις πληροφορίες που συλλέχτηκαν έτσι ώστε να την ευρετηριάσει. Ο spider είναι ένα αυτόματο πρόγραμμα που επισκέπτεται μια δεδομένη ιστοσελίδα και ανακτά το περιεχόμενο της εξάγοντας τον HTML κώδικα της. Η διαδικασία ανάγνωσης της ιστοσελίδας και εξαγωγή πληροφοριών ονομάζεται ανίχνευση ή crawling.

Η διαδικασία ανίχνευσης ορίζεται ως εξής. Ο spider μιας μηχανής αναζήτησης μεταφορτώνει μια σελίδα και την αποθηκεύει στον διακομιστή της μηχανής αναζήτησης, καθώς και ένα δεύτερο πρόγραμμα, γνωστό ως indexer, εξάγει διάφορες χρήσιμες πληροφορίες σχετικά με την σελίδα, όπως οι λέξεις που περιέχει, η βαρύτητα τους καθώς και οι υπερσυνδέσεις που περιέχει η σελίδα.

Οι ιδιοκτήτες ιστοσελίδων άρχισαν να συνειδητοποιούν πόσο σημαντικό είναι για τις ιστοσελίδες τους να εμφανίζονται ψηλά στην λίστα των αποτελεσμάτων μιας αναζήτησης. Το γεγονός οδήγησε στην δημιουργία καλών αλλά και κακόβουλων πρακτικών για την βελτιστοποίηση της θέσης τους. Σύμφωνα με τον αναλυτή Danny Sullivan, η φράση «*search engine optimization*» χρησιμοποιήθηκε πρώτη φορά το 1997.

Παλαιότερες εκδόσεις των αλγορίθμων αναζήτησης στηρίζονταν στις πληροφορίες που έδινε ένας webmaster κατά την εκχώρηση της ιστοσελίδας του στον κατάλογο της μηχανής αναζήτησης, όπως είναι λέξεις-κλειδιά, αλλά και στις πληροφορίες που αντλούσαν μέσω των μεταδεδομένων. Τα μεταδεδομένα αποτελούν έναν οδηγό για τα περιεχόμενα μιας σελίδας. Η χρήση των μεταδεδομένων για την καταχώρηση και καταγραφή των στοιχείων μιας ιστοσελίδας αποδείχθηκε αναξιόπιστος τρόπος διότι πολλές από τις λέξεις-κλειδιά που δίδονταν ήταν ανακριβείς και ορισμένες φορές άσχετες με το πραγματικό περιεχόμενο. Ανακριβή, ψευδή και ελλιπή μεταδεδομένα προκαλούσαν σύγχυση στην ευρετηρίαση και την κατάταξη των ιστοσελίδων διότι εμφανίζονταν σελίδες με άσχετο, όσο αφορά τον όρο αναζήτησης του χρήστη, περιεχόμενο. Επίσης, υπήρχε ένας τρόπος χειραγώγησης της κατάταξης. Η αλλαγή μερικών χαρακτηριστικών στον HTML κώδικα μπορούσε να ξεγελάσει τις μηχανές αναζήτησης

και οι ιστοσελίδες είχαν την δυνατότητα να καταταχθούν υψηλότερα από την πραγματική τους αξία.

Οι μηχανές αναζήτησης υπέφεραν από την κατάχρηση γιατί βασίζονταν σε απλές μεθόδους αξιολόγησης, όπως η πυκνότητα των λέξεων-κλειδιών, μια μέθοδο που ήταν εύκολη στην χειραγώγηση της με την επανάληψη των λέξεων-κλειδιών ως κείμενο σε μια σελίδα. Λαμβάνοντας υπ όψιν τις καταχρήσεις, οι μηχανές αναζήτησης αμύνθηκαν με την δημιουργία πολυπλοκότερων αλγορίθμων αξιολόγησης χρησιμοποιώντας παράγοντες που δεν μπορεί να χειραγωγήσει εύκολα κάποιος χρήστης.

Οι απόφοιτοι του Πανεπιστημίου του Stanford , Larry Page και Sergey Brin, ανέπτυξαν ένα σύστημα, μια μηχανή αναζήτησης, που στηρίζεται σε μαθηματικό αλγόριθμο για την αξιολόγηση των σελίδων. Ο αριθμός που υπολογίζεται από τον αλγόριθμο ονομάστηκε Page Rank και βασίζεται κυρίως στις εισερχόμενες συνδέσεις από άλλες ιστοσελίδες... (βλ Page Rank).

Ο Page και ο Brin ίδρυσαν την Google το 1998. Η Google προσέλκυσε πολλούς χρήστες λόγω της απλότητας του σχεδιασμού αλλά και της λειτουργίας της. Οι παράγοντες όπως το Page Rank, η ανάλυση υπερσυνδέσεων, τα μεταδεδομένα καθώς και άλλα στοιχεία μιας σελίδας χρησιμοποιήθηκαν από την Google προς αποφυγή της χειραγώγησης της αξιολόγησης μιας ιστοσελίδας. Αν και το Page Rank ήταν δύσκολο να χειραγωγηθεί, οι webmasters ανέπτυξαν τρόπους και εργαλεία που θα αύξαναν τον αριθμό των εσωτερικών συνδέσεων στην σελίδα τους, όπως ανταλλαγή, αγοραπωλησία μαζικών συνδέσμων καθώς και φάρμες συνδέσμων (link farms)

Η ανταλλαγή συνδέσμων ήταν απλή διαδικασία και εφαρμόζεται μέχρι σήμερα. Ο ιδιοκτήτης μιας ιστοσελίδας ζητά να τοποθετηθεί σύνδεσμος ως προς την σελίδα του από έναν άλλο ιδιοκτήτη και να πράξει και εκείνος το ίδιο. Με αυτό τον τρόπο οι ιστοσελίδες ανταλλάζουν συνδέσμους.

Η αγοραπωλησία είναι παρόμοιος τρόπος με την ανταλλαγή συνδέσμων με την διαφορά ότι ο ιδιοκτήτης μιας ιστοσελίδας πληρώνει τον ιδιοκτήτη μιας άλλης έτσι ώστε να εισάγει έναν σύνδεσμο του. Αυτό δημιούργησε ένα εμπόριο μαύρου χρήματος και ανεξέλεγκτης αγοραπωλησίας.

Οι φάρμες συνδέσμων (link farms) είναι ιστοσελίδες χωρίς νόημα και περιεχόμενο. Ο σκοπός τους είναι η δημοσίευση όλο και περισσότερων συνδέσμων προς άλλες ιστοσελίδες. Κάθε ιδιοκτήτης ιστοσελίδας μπορεί να προσθέσει τον σύνδεσμο του χωρίς κόστος ή έγκριση από τον ιδιοκτήτη της φάρμας.

Ο σκοπός των παραπάνω τεχνασμάτων ήταν η αύξηση των εσωτερικών συνδέσεων προς μια ιστοσελίδα, πράγμα που σήμαινε την καλύτερη αξιολόγηση τους όσο αφορά το PR. Ο μοναδικός σκοπός τους ήταν το link spamming.

Το 2004, οι μηχανές αναζήτησης ενσωμάτωσαν ένα εύρος από μυστικούς παράγοντες στους αλγόριθμους κατάταξης ώστε να καταπολεμήσουν την χειραγώγηση των συνδέσμων. Η Google δηλώνει ότι κατατάσσει τις ιστοσελίδες λαμβάνοντας υπόψη πάνω από 200 διαφορετικά στοιχεία. Οι τρεις μεγαλύτερες μηχανές αναζήτησης, Google, Yahoo και η Bing της Microsoft, δεν αποκαλύπτουν τον αλγόριθμο που χρησιμοποιούν. Αξιοσημείωτοι αναλυτές βελτιστοποίησης ιστοσελίδων, όπως ο Rand Fishkin, Barry Schwartz, Aaron Wall και Jill Whalen, μελέτησαν την βελτιστοποίηση ιστοσελίδων από διαφορετικές σκοπιές και δημοσίευσαν τις έρευνες και τα πορίσματα τους σε περιοχές συζητήσεων και ιστολόγια.

Το 2005 η Google άρχισε σταδιακά να αλλάζει τον τρόπο κατάταξης των αποτελεσμάτων. Ο νέος τρόπος κατάταξης ήταν η προσαρμογή τους ανάλογα με τον χρήστη, χρησιμοποιώντας το ιστορικό αναζητήσεων του. Το 2008 ο Bruce Clay δήλωσε ότι ο αγώνας για την καλύτερη κατάταξη μπορεί να σταματήσει λόγω της προσωπικής παραμετροποίησης των αποτελεσμάτων. Η ενασχόληση με το Page Rank θα είναι ανούσια διότι τα αποτελέσματα αναζήτησης θα διαφέρουν από χρήστη σε χρήστη.

Το 2007 η Google ξεκίνησε καμπάνια ενάντια στην αγοραπωλησία συνδέσεων που επηρέαζαν το Page Rank και το 2009 πήρε μέτρα να μετριάσει την μη θεμιτή χειραγώγηση του με την χρήση του HTML χαρακτηριστικού «nofollow» στις υπερσυνδέσεις.

### **2.1.10 Βελτιστοποίηση Σελίδων και Κοινωνικά Δίκτυα**

Τα κοινωνικά δίκτυα είναι η επανάσταση στον χώρο του διαδικτύου αλλά και του τρόπου ζωής στην τρέχουσα χρονική περίοδο. Έχουν φέρει αλλαγές σε επίπεδο

δο επικοινωνίας και σχέσεων μεταξύ των ανθρώπων, απασχόλησης, ψυχαγωγίας, ενημέρωσης, ασφάλειας και άλλους τομείς που έχουν βλέψεις εκτός των ορίων της πτυχιακής εργασίας. Ένας από τους τομείς που απασχολεί τις ιστοσελίδες είναι η προώθηση τους μέσω αυτών.

Τα κοινωνικά δίκτυα δεν είναι όλα ίδια. Μερικά προσφέρουν υπηρεσίες γνωριμιών, προφίλ αλλά και δικτύωσης όπως είναι το Facebook ενώ άλλα δίκτυα, όπως το digg, υπηρεσίες διαμοιρασμού ενδιαφερόντων συνδέσμων. Κάθε τύπος δικτύου, έχει δικό του ρόλο. Όσο αφορά την βελτιστοποίηση σελίδων μεγαλύτερο ενδιαφέρον εκδηλώνεται στα δίκτυα κοινωνικής προσθήκης/διακίνησης περιεχομένου – social bookmarking networks, όπως το digg.

Η εκμετάλλευση αυτών των δικτύων γίνεται με το χτίσιμο των υπερσυνδέσεων ως προς μια ιστοσελίδα (backlinks). Τα κοινωνικά δίκτυα έχουν υψηλό PR, γεγονός που μπορεί να απογειώσει το PR μιας ιστοσελίδας που θα αναρτηθεί σε αυτά. Τα δίκτυα θυμίζουν αρκετές φορές τις παλαιότερες φάρμες συνδέσμων, με την διαφορά ότι το περιεχόμενο χτίζεται μόνο από ανθρώπους. Θα πρέπει να γίνεται σταδιακά προσθήκη συνδέσμων περιεχομένου από διαφορετικές ιστοσελίδες και σε μη τακτά χρονικά διαστήματα, ώστε να μην θεωρηθεί ως spammer ο αποστολέας.

Υπάρχουν πολλά κοινωνικά δίκτυα, αλλά δεν είναι όλα αποτελεσματικά. Πολλά κοινωνικά δίκτυα χρησιμοποιούν την τεχνική του «nofollow» στις υπερσυνδέσεις. Το Facebook και το Twitter για παράδειγμα το χρησιμοποιεί και δεν δίνει αξία στις σελίδες οι οποίες δημοσιεύονται εκεί. Αυτό σημαίνει ότι η σελίδα του χρήστη δεν θα λάβει κάποια θετική ψήφο για την αξιολόγηση της. Το digg σε αντίθεση έχει υψηλό PR και δίνει θετική ψήφο σε μια ιστοσελίδα. Αρκεί όλα να γίνονται με μέτρο. Τα πιο γνωστά κοινωνικά δίκτυα με *DoFollow* συνδέσεις είναι το Slash Dot και το Digg.

### **2.1.11 Βελτιστοποίηση Δυναμικών ιστοσελίδων**

Οι δυναμικές ιστοσελίδες, όπως είναι οι ιστοσελίδες σχεδιασμένες σε Flash ή προγραμματισμένες με μια δυναμική γλώσσα προγραμματισμού, είναι κυρίως οι σελίδες που αντιμετωπίζουν προβλήματα με την βελτιστοποίησή τους, γι αυτό και χρειάζονται ιδιαίτερη αντιμετώπιση.

## **Flash**

Οι ιστοσελίδες σχεδιασμένες σε Flash παρουσιάζουν υψηλό γραφικό φρόνημα αλλά δεν βοηθούνται καθόλου από τις μηχανές αναζήτησης. Μια Flash ιστοσελίδα προκαλεί μεγάλη εντύπωση στον επισκέπτη, κυρίως όταν είναι μοναδικός ο τρόπος της κίνησης και των γραφικών. Ο κώδικας τους όμως δεν είναι τίποτα παραπάνω από ένα αντικείμενο σε HTML το οποίο δεν μπορεί να διαβαστεί από κάποιο spider ή άλλο πρόγραμμα αποκωδικοποίησης. Ως αποτέλεσμα, οι μηχανές αναζήτησης δεν μπορούν να διαβάσουν το περιεχόμενο της ιστοσελίδας και δεν την καταχωρούν σωστά στις βάσεις δεδομένων τους, παρά μόνο την αρχική σελίδα με ότι πληροφορίες μπορούν να αποκομίσουν, όπως ο τίτλος.

Λύση στο θέμα των flash ιστοσελίδων είναι η δημιουργία μιας ιστοσελίδας σε μορφή html με τα χαρακτηριστικά και κείμενα που έχει η flash έκδοση της έτσι ώστε η μηχανή αναζήτησης να μπορεί να διαβάζει το κείμενο και να κάνει την ευρετηρίαση καλύτερα. Η πρακτική αυτή ωφελεί και τους επισκέπτες της σελίδας που δεν έχουν γρήγορες συνδέσεις ή συνδέονται από φορητές συσκευές όπως PDA, κινητά και άλλες.

## **Ajax**

Πρόβλημα αντιμετωπίζουν εξίσου και οι δυναμικές ιστοσελίδες που χρησιμοποιούν τεχνολογία AJAX. Όπως είναι γνωστό, οι spiders δεν μπορούν να ακολουθούν ή να εκτελούν κώδικα JavaScript για λόγους ασφαλείας. Η τεχνολογία AJAX (*Asynchronous JavaScript And XML*) είναι βιβλιοθήκη της JavaScript και η βελτιστοποίηση της είναι σχετικά δύσκολη και επίπονη διαδικασία παρόλο την καλαισθησία και την διευκόλυνση που παρέχει στον χρήστη, όπως η δυναμική αλλαγή του περιεχομένου.

Ένας τρόπος αντιμετώπισης του προβλήματος, κυρίως για τους συνδέσμους με JavaScript, είναι η χρήση της εντολής *return false*. Για παράδειγμα, αν υπάρχει ένας σύνδεσμος που φορτώνει τα περιεχόμενα μιας σελίδας μέσα σε ένα div, εκτός από την κλήση της ρουτίνας που γίνεται στα head tags για την μέθοδο δυναμικής φόρτωσης, δηλώνεται ως εξής.

```
<a href="javascript:ajaxpage('test.htm', 'contentarea');">test</a>
```

Ο παραπάνω σύνδεσμος όμως περιέχει Java Script και η μηχανή αναζήτησης θα τον αγνοήσει. Ο σκοπός όμως των συνδέσμων είναι να ακολουθούνται από τις μηχανές αναζήτησης. Για να βελτιστοποιηθεί ο σύνδεσμος, πρέπει να τροποποιηθεί σε:

```
<a href="test.htm" onclick="javascript:ajaxpage('test.htm', 'contentarea');  
return false">test</a>
```

Η εντολή *return false* είναι ένα γεγονός της Java Script που ακυρώνει την προκαθορισμένη συμπεριφορά του γεγονότος που το πυροδότησε. Δηλαδή, ο σύνδεσμος καλεί την σελίδα test.htm αλλά ο Java Script κώδικας δεν του επιτρέπει να τον καλέσει. Ως ακολούθως, συνεχίζει στην εκτέλεση του JavaScript κώδικα και φορτώνει το περιεχόμενο της σελίδας στο div που δηλώθηκε. Ο spider όμως θα αγνοήσει τον JavaScript και θα ακολουθήσει τον σύνδεσμο κανονικά και θα φορτώσει την σελίδα test.htm σε ξεχωριστή σελίδα. Αν και αισθητικά δεν είναι το σωστό, γιατί η σελίδα test.htm θα περιέχει μόνο κείμενο και όχι τα γραφικά της σελίδας, η μηχανή αναζήτησης αντλεί τα στοιχεία που χρειάζεται από την σελίδα.

### **Sessions IDs**

Τέλος, πρόβλημα αντιμετωπίζουν ιστοσελίδες που χρησιμοποιούν session id. Το session id είναι ένας μοναδικός αριθμός που συνήθως τοποθετείται σε ένα cookie ή στην διεύθυνση της σελίδας ως παράμετρο που βοηθά στην συλλογή δεδομένων για τις κινήσεις ενός επισκέπτη. Ένα παράδειγμα διεύθυνσης με session id είναι το [selida.gr/index.php?sid=6543df1Cujmeud83ebe894e5](http://selida.gr/index.php?sid=6543df1Cujmeud83ebe894e5)

Το κυρίως πρόβλημα με τα Session IDs (SIDs) είναι ότι οι spider δεν τα ευρετηριάζουν σωστά. Όταν μια μηχανή αναζήτησης ευρετηριάζει μια ιστοσελίδα στέλνει περισσότερα από ένα spider για την ανάκτηση πληροφοριών. Σύμφωνα με τον μηχανισμό SID, κάθε spider παίρνει διαφορετικό id και πιστεύει ότι οι διευθύνσεις που επισκέπτεται είναι μοναδικές. Την ίδια άποψη έχουν και οι άλλοι spiders, με την διαφορά ότι δέχονται διαφορετικές διευθύνσεις για τις ίδιες σελίδες λόγω του διαφορετικού SID. Ως ακολούθως, στις επόμενες επισκέψεις οι spiders θα λάβουν διαφορετικό id και θα αντιληφθούν ότι είναι άλλες σελίδες και θα τις ευρετηριάσουν και αυτές. Αν γίνει αυτό, μια σελίδα μπορεί να ευρετηριαστεί πάρα πολλές φορές (λόγω των πολλών συνδυασμών σε 32bit χαρακτήρες



στην ανάθεση sid) γεγονός που δημιουργεί σύγχυση και επιβαρύνει τις μηχανές και καταστρέφει το ευρετήριο τους. Ως μέτρο αντιμετώπισης του φαινομένου της πολλαπλής ευρετηρίασης από τις μηχανές αναζήτησης είναι η αγνόηση διευθύνσεων που περιέχουν παραμέτρους όπως "sid=", id=".

Μοναδική λύση στο πρόβλημα των SIDs είναι μη χρήση τους. Δυστυχώς η βελτιστοποίηση μιας ιστοσελίδας περιορίζει αρκετές εικαστικές και τεχνικές λεπτομέρειες καθώς και ευκολίες προγραμματισμού. Τα session ids μπορούν κάλλιστα να αντικατασταθούν με άλλες μεθόδους όπως η χρήση cookies, μεταβλητές στον server ή αποθήκευση των SIDs στην βάση δεδομένων ώστε να επιτευχτεί η αποθήκευση προσωρινών δεδομένων των χρηστών.

## 2.2 Στοιχεία και Τεχνικές

### 2.2.1 Domain

Η διεύθυνση μιας ιστοσελίδας παίζει πλέον πολύ σημαντικό ρόλο στην βελτιστοποίηση της, σε αντίθεση με παλαιότερα όπου το λεγόμενο URL (*Uniform Resource Locator - Ενιαίος Εντοπιστής Πόρων*), δεν έπαιζε μεγάλο ρόλο στην αξιολόγηση της. Στόχος της ιστοσελίδας ήταν η εύρεση και χρήση μιας μικρής και εύκολης, στην απομνημόνευση, διεύθυνσης χωρίς ο χρήστης να χρειάζεται να πληκτρολογεί πολλούς χαρακτήρες, πόσο μάλλον να θυμάται ειδικά σύμβολα όπως είναι η παύλα ή η κάτω παύλα «-,\_»

Με την αλλαγή των αλγορίθμων αναζήτησης και την πάροδο των χρόνων, το domain αποτελεί πλέον μείζον παράγοντα για την βελτιστοποίηση μιας ιστοσελίδας. Κάθε ιστοσελίδα θα πρέπει να περιέχει μια ή δυο λέξεις-κλειδιά, με ή χωρίς χρήση ειδικών χαρακτήρων, στην διεύθυνση της ώστε οι μηχανές αναζήτησης να εντοπίζουν εύκολα το περιεχόμενο της και να θεωρούν ότι η σελίδα είναι πολύ σχετική με το αντικείμενο. Πλέον, δεν παίζει μεγάλο ρόλο η έκταση του αλλά οι λέξεις που περιέχει.

Επίσης, σημαντικό ρόλο στην βελτιστοποίηση παίζει η κατάληξη του domain. Ο ρόλος της κατάληξης ενός top-level domain αφορά κυρίως την γεωγραφική βελτιστοποίηση της ιστοσελίδας. Η γεωγραφική βελτιστοποίηση δεν είναι τίποτα παραπάνω από τον γεωγραφικό προσδιορισμό του κοινού μιας ιστοσελίδας και της εθνικότητας του. Για παράδειγμα, ιστοσελίδες που διαθέτουν .gr

domain σχετικά με τους όρους αναζήτησης σε ελληνικό κοινό θα εμφανιστούν πρώτες στην λίστα των αποτελεσμάτων ενώ αντίθετα μπορεί να μην εμφανιστούν στην αναζήτηση ενός αγγλόφωνου χρήστη. Όταν μια σελίδα απευθύνεται στο παγκόσμιο κοινό συνιστάται η χρήση της κατάληξης .com

Τέλος, η ηλικία ενός domain παίζει ρόλο στην βελτιστοποίηση του. Όσο πιο παλιό είναι, τόσο ποιοτικότερο θεωρείται το υλικό που περιέχει σχετικά με το θέμα της σελίδας, μιας που θα δαπανηθεί μεγαλύτερο ποσό για την αγορά του, αλλά και τόσες περισσότερες συνδέσεις προς την σελίδα αυτή θα υπάρχουν από άλλους χρήστες. Για παράδειγμα, η αγορά ενός domain ηλικίας 2χρόνων μπορεί να έχει PR 3 ενώ ένα νέο domain ξεκινάει με PR 0. Υπάρχουν ειδικές υπηρεσίες στις οποίες αγοράπωλούνται domains ανάλογα την ηλικία δειγματίζοντας πόσοι σύνδεσμοι δείχνουν σε αυτό το domain (inbound links) ή το PR τους.

Για μια επιχείρηση η λέξη-κλειδί *business* είναι ζωτικής σημασίας γιατί όσοι χρήστες αναζητούν κάποιο επιχειρηματικό όρο ή εταιρία χρησιμοποιούν και αυτή την λέξη στους όρους αναζήτησης. Σύμφωνα με την CNET News το domain *business.com* πωλήθηκε το 1999 έναντι \$7,5 εκατομμυρίων<sup>2</sup> ενώ το σημερινό κόστος του ανέρχεται στα \$400 εκατομμύρια. Το κόστος ενός τόσο σπουδαίου domain δείχνει πόσο σημαντικό ρόλο μπορεί να παίζει στην βελτιστοποίηση ιστοσελίδων.

Η χρήση δύο ή περισσότερων λέξεων στο όνομα του domain με την κατάλληλη δομή της σελίδας σύμφωνα με τις λέξεις-κλειδιά αποτελεί ένα δυνατό όπλο στον ανταγωνισμό. Αν μια ιστοσελίδα έχει ένα όνομα domain που ταιριάζει με τις λέξεις-κλειδιά της, έχει ήδη βελτιστοποιηθεί κατά 25% με στόχο τις υψηλότερες θέσεις στα αποτελέσματα των μηχανών αναζήτησης. Παρατηρώντας ότι μόνο η χρήση του σωστού ονόματος αποτελεί το ¼ της σωστής βελτιστοποίησης, διαπιστώνεται ότι πρέπει να υπάρξει συζήτηση και σκέψη προτού αγοραστεί κάποιο domain.

Όμως, μεγάλη προσοχή θα πρέπει να δοθεί στον τρόπο που θα χρησιμοποιηθούν οι λέξεις-κλειδιά. Το όνομα του domain θα πρέπει να είναι σωστά δομημένο και να αποτελεί μια έκφραση στον γραπτό λόγο που περιέχει κάποιο

---

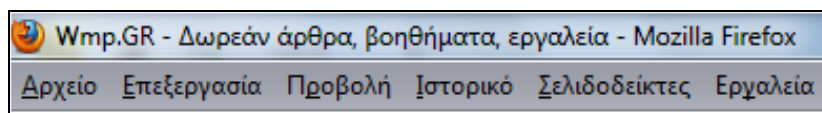
<sup>2</sup> MSNBC, *Domain name sells for \$2.75 million*, Ιούλιος 2004

νόημα και όχι η απλή εισαγωγή των λέξεων κλειδιών. Η χρήση ενός ονόματος όπως lose-weight-fast-cheap.com είναι μια κακή επιλογή για ιστοσελίδα που εμπορεύεται υπηρεσίες αδυνατίσματος και δεν θα φέρει καθόλου τα επιθυμητά αποτελέσματα.

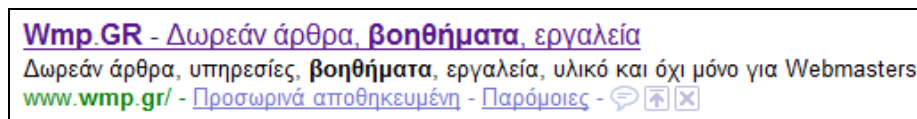
Μια ιστοσελίδα που αναφέρεται στην πώληση ανθέων-φυτών στην Ελλάδα καλύπτεται πλήρως με ένα domain σχετικό με το εμπόρευμα του, λόγω του ότι είναι οι κυριότερες λέξεις-κλειδιά γι αυτό. Για την σωστή βελτιστοποίηση του domain προτιμάται το *www.flowershop.gr* παρά η επωνυμία της εταιρίας *www.selimidis.gr*

### 2.2.2 Τίτλος Σελίδας

Η χρήση τίτλων σε μια ιστοσελίδα, εκτός της αισθητικής βελτίωση στο περιβάλλον του χρήστη, συμβάλει και στην βελτίωση κατάταξης της. Προτού γίνει αναφορά στο κομμάτι της βελτιστοποίησης του τίτλου, πρέπει να αναφερθεί ότι ο τίτλος, ή οι πρώτοι 60 με 70 χαρακτήρες του όταν τους υπερβαίνει, αποτελεί το όνομα της ιστοσελίδας που εμφανίζεται στα αποτελέσματα αναζήτησης μιας μηχανής αναζήτησης όπως φαίνεται στις ακόλουθες εικόνες (Εικόνα 2.2 και Εικόνα 2.3). Σύμφωνα με τα παραπάνω, θα πρέπει να είναι εμφανίσιμος, περιεκτικός και μετριασμένος σε μέγεθος.



Εικόνα 2.2: Τίτλος σελίδας σε φυλλομετρητή



Εικόνα 2.3: Τίτλος σελίδας στο Google

Για την σωστή βελτιστοποίηση, ο τίτλος μιας ιστοσελίδας θα πρέπει να περιέχει μια μικρή περιγραφή του περιεχομένου της. Θα πρέπει να περιέχει μερικές κύριες λέξει-κλειδιά ή φράσεις ώστε να επηρεάζονται θετικά οι μηχανές

αναζήτησης. Επίσης, ο τίτλος της ιστοσελίδας δεν πρέπει να είναι στατικός αλλά να μεταβάλλεται γιατί κάθε σελίδα είναι μοναδική σε περιεχόμενο. Μια καλή πρακτική, κυρίως για την αρχική σελίδα, είναι η χρήση μιας πολύ συνοπτικής, αλλά περιεκτικής σε λέξεις-κλειδιά, περιγραφής της σελίδας και του ονόματος της ιστοσελίδας χωρίς να ξεπερνά τις 60 λέξεις.

Προσοχή θα πρέπει να δοθεί και στο ύφος του τίτλου. Οι μηχανές αναζήτησης ελέγχουν τους τίτλους αν έχουν κάποιο ύφος όπως ερώτηση, υπόσχεση και άλλα σύμφωνα με το ειδικό λεξιλόγιο της γλώσσας της ιστοσελίδας. Για παράδειγμα, η χρήση της φράσης «πώς να» αναφέρεται κυρίως σε ιστοσελίδες που παρέχουν πληροφορίες για το πως θα κάνει ο χρήστης κάτι συγκριμένο. Συνήθως η χρήση ερωτηματικών λέξεων παραπέμπει σε σελίδες όπως *fora*, *helpdesk* και άλλες.

Μια κακή πρακτική που δεν συνιστάται, είναι η πληκτρολόγηση των λέξεων κλειδιών στον τίτλο μιας ιστοσελίδας, δηλαδή ό,τι περιέχει το *meta tag keywords*. Η πρακτική αυτή έχει διαφορά γιατί οι λέξεις-κλειδιά δεν χρησιμοποιούνται για την δόμηση του τίτλου αλλά για την απλή αναφορά τους. Αυτό αποτελεί το λεγόμενο *keyword spamming* (άσκοπη χρήση των λέξεων-κλειδιών) και αναλυτές βελτιστοποίησης υποστηρίζουν ότι είναι λάθος τρόπος χρήσης της τεχνικής του τίτλου και δεν φέρει τα επιθυμητά αποτελέσματα.

Για μια ιστοσελίδα με τα άνθη-φυτά ένας βέλτιστος τίτλος θα μπορούσε να είναι «Άνθη, φυτά, δώρα σε όλη την Ελλάδα | Flowershop.GR». Κοινές λέξεις όπως το «και», «η» παραβλέπονται από τις μηχανές αναζήτησης.

### **2.2.3 Περιγραφή Ιστοσελίδας**

Τα μεταδεδομένα είναι ειδικές οδηγίες που τοποθετούνται στην αρχή της σελίδας και περιέχουν χρήσιμες πληροφορίες για τις μηχανές αναζήτησης. Λόγω κατάχρησης όμως δεν βοηθούν ενεργά στην βελτιστοποίηση μιας σελίδας αλλά οι μηχανές αναζήτησης τα χρησιμοποιούν για να πάρουν διάφορες πληροφορίες για την ιστοσελίδα.

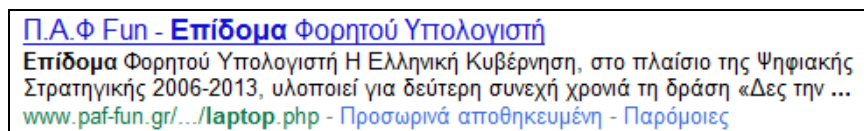
Παρόλο που τα περισσότερα μεταδεδομένα έχασαν την αξία τους, ένα μόνο *tag* εξακολουθεί να επηρεάζει θετικά μια σελίδα. Το μεταδεδομένο της περιγραφής (*meta description*) χρησιμοποιείται από τις μηχανές αναζήτησης ως

περιγραφή της σελίδας για την εμφάνιση της στα αποτελέσματα αναζήτησης. Εναλλακτικά, η μηχανή αναζήτησης μπορεί να εμφανίσει ένα απόσπασμα κειμένου από την σελίδα ως περιγραφή, θεωρώντας το πιο σημαντικό για τον όρο αναζήτησης του χρήστη και αγνοώντας εντελώς το meta description (βλ. Εικόνα 2.5)

Καλή πρακτική αποτελεί η ακριβής περιγραφή της σελίδας στο meta description. Η σωστή περιγραφή πρέπει να είναι πλήρης, δηλαδή να περιέχει όλες τις πληροφορίες που μπορεί να βρει κάποιος στην σελίδα, συνοπτική και ακριβής. Δεν θα πρέπει να περιέχει γενικούς όρους όπως «η σελίδα μιλάει για...» αλλά μια λακωνική περιγραφή, δομημένη κυρίως με λέξεις-κλειδιά. Τέλος κάθε σελίδα θα πρέπει να περιέχει την δική της, μοναδική περιγραφή.



Εικόνα 2.4: Απόσπασμα σελίδας



Εικόνα 2.5: Αποτέλεσμα περιγραφής Google

## 2.2.4 Ετικέτες Επικεφαλίδων

Ένα από τα σημεία που αγνοούν οι περισσότεροι προγραμματιστές ιστοσελίδων είναι η χρήση των επικεφαλίδων (h1, h2,...h6 tags) σε μια σελίδα, ενώ πολλοί χρησιμοποιούν γραφικά για επικεφαλίδες σε ζωτικά σημεία της σελίδας. Βάσει κριτηρίων βελτιστοποίησης, οι μηχανές αναζήτησης αντιλαμβάνονται το περιεχόμενο μιας σελίδας με την περίληψη των επικεφαλίδων της και θέτουν περισσότερο βάρος σε αυτές παρά στο σώμα κειμένου. Η χρήση στοχευμένων λέξεων ή φράσεων στις επικεφαλίδες, και κυρίως στην h1, αξιολογείται πολύ θετικά από μια μηχανή αναζήτησης.

Ένας ακόμη παράγοντας είναι η ιεραρχία τους. Αφού οι μηχανές αναζήτησης δίνουν μεγάλο βάρος σε αυτή, η επικεφαλίδα 1<sup>ου</sup> επιπέδου, h1, θα πρέπει να περιέχει κείμενο με την κυριότερη λέξη-κλειδί της ιστοσελίδας. Οι επικεφαλίδες 2<sup>ου</sup> θα περιέχουν με την σειρά τους τις επόμενες σημαντικές λέξεις-κλειδιά και ούτω καθεξής. Αφού δημιουργηθεί μια σωστή ιεραρχία, με την ανάγνωση μόνο των επικεφαλίδων θα πρέπει να αντιλαμβάνεται ο χρήστης το περιεχόμενο της σελίδας.

Σε ορισμένες περιπτώσεις οι επικεφαλίδες δεν ταιριάζουν με την αισθητική της ιστοσελίδας, είτε για το μέγεθος τους, είτε για την μορφή τους. Σε αυτές τις περιπτώσεις γίνεται εκμετάλλευση των CSS όπου δίνεται οποιαδήποτε μορφή απλού κειμένου σε επικεφαλίδες. Τα αποτελέσματα της χρήσης των CSS είναι η αισθητική βελτίωση των επικεφαλίδων στον αναγνώστη και οι μηχανές αναζήτησης θα τις διαβάζουν κανονικά μέσα στον HTML κώδικα. Τεχνικές όπως μορφοποίηση της επικεφαλίδας ως απλό κείμενο και εισαγωγή όλου του περιεχομένου της σελίδας ανάμεσα σε <h1></h1> tags δεν λειτουργούν.

Για το παράδειγμα του καταστήματος με τα άνθη, οι χρήσεις ως <h1>Άνθη, Φυτά, Δώρα</h1> και <h2>Ηλεκτρονικό κατάστημα</h2>, <h2>Δώρα</h2>, <h2>Ποικηλίες ανθέων</h2>, <h2>Διανομές</h2> θα έδιναν τα πλήρη χαρακτηριστικά της σελίδας στην περίληψη βάσει επικεφαλίδων.

## 2.2.5 Περίληψη Σελίδας

Μια δύσκολη με την εφαρμογή αλλά αποτελεσματική τεχνική είναι η χρήση του σημαντικότερου περιεχομένου μιας ιστοσελίδας στην αρχή της. Ένας τρόπος ανάλυσης μιας ιστοσελίδας από τις μηχανές αναζήτησης είναι η ανάγνωση των 25 πρώτων λέξεων που θα βρεθούν στην αρχή του HTML Body. Σύμφωνα με το παραπάνω κριτήριο, είναι θεμιτό να παρουσιάζεται το κύριο περιεχόμενο της ιστοσελίδας εκεί. Το σημαντικότερο λάθος που παρατηρείται σε αυτό το κριτήριο είναι η χρήση ενός μενού για την σελίδα. Λέξεις που συναντώνται συχνά, όπως «Αρχική σελίδα», «Επικοινωνία», είναι άσχετες με το περιεχόμενο που πρέπει να προβληθεί σε μια μηχανή αναζήτησης.

Υπάρχουν διάφορα εργαλεία ανάλυσης στο Διαδίκτυο όπου ο χρήστης μπορεί να δει τι διαβάζει μια μηχανή αναζήτησης από την σελίδα του. Συνηθέσ-

τερο φαινόμενο είναι η χρήση του ονόματος της ιστοσελίδας και ένα μικρό κείμενο-περιγραφή για το τι έχει να προσφέρει, χρησιμοποιώντας πάντα λέξεις-κλειδιά. Ακόμη και ένα σχετικό μενού με το ουσιαστικό περιεχόμενο της ιστοσελίδας φέρει θετικότερα αποτελέσματα από ένα γενικό μενού.

Το πρόβλημα συναντάται κυρίως στον τρόπο δόμησης της σελίδας. Η χρήση των `divs` ή των πινάκων με την παραδοσιακή φορά από πάνω προς τα κάτω και αριστερά προς τα δεξιά μεταφέρει το κυρίως περιεχόμενο αρκετά κάτω από τις 25λέξεις. Υπάρχουν τεχνικές που υλοποιούνται με την χρήση της CSS για την δόμηση των πολύστηλων layout ώστε να εμφανίζονται στον κώδικα την κεντρική στήλη ψηλά.

Για το παράδειγμα των φυτών, η χρήση μιας μικρής αναφοράς των υπηρεσιών, όπως στο παραπάνω κριτήριο, με την σωστή σύνταξη και αισθητική τοποθέτηση, αποτελεί μια καλή λύση.

### 2.2.6 Έντονες/Πλάγιες Λέξεις

Ένα ακόμη κριτήριο σωστής μορφοποίησης αλλά και βελτιστοποίησης μιας ιστοσελίδας είναι η χρήση έντονης γραφής (`<b>`, `<strong>`) καθώς και η πλάγια γραφή (`<i>`, `<em>`) Ένας spider μιας μηχανής αναζήτησης κάθε φορά που επισκέπτεται μια σελίδα διαβάζει όλο το περιεχόμενο της. Όταν συναντά λέξεις με έντονη ή πλάγια γραφή, τις θεωρεί σημαντικό, για την σελίδα, περιεχόμενο και τις δίνει ιδιαίτερη προσοχή.

Άλλοτε τα *strong* και *em* tags δεν είχαν την ίδια βαρύτητα με τα αντίστοιχα τους *b* και *i* αλλά θεωρούνταν κατώτερα. Η διάκριση όμως επηρεάζει αρνητικά τις ιστοσελίδες που επιθυμούν την μέγιστη βελτιστοποίηση αλλά και την συμβατότητα με τους φυλλομετρητές για ΑμΕΑ διότι οι τελευταίοι παραβλέπουν όλες τις λέξεις που δηλώνονται με bold και italics, γιατί τα συγκεκριμένα tags προσδίδουν μόνο μορφοποίηση. Αργότερα ανακοινώθηκε από την Google ότι ο σύγχρονος αλγόριθμος έχει ίδια συμπεριφορά απέναντι στα *b,i* και *strong,em* αντίστοιχα.

Δεν πρέπει να γίνεται κατάχρηση αυτής της τεχνικής. Η χρήση της μορφοποίησης σε ολόκληρη την σελίδα ή σε παραγράφους απλά αναιρεί το κριτήριο και αγνοείται το κείμενο, όσο αφορά τον έλεγχο σημαντικότητας του. Επί-

σης, καλό είναι πρώτα να λαμβάνονται υπ όψιν οι επισκέπτες της σελίδας και μετά η βελτιστοποίηση. Το κείμενο θα πρέπει να είναι ευανάγνωστο και να μην κουράζεται ο αναγνώστης με την άσκοπη εναλλαγή μεταξύ της κανονικής, της έντονης και της πλάγιας γραφής. Η σωστή μορφοποίηση, όπως έντονη γραφή για τονισμό και έμφαση, και πλάγια γραφή για επεξήγηση, μαζί με την χρήση των λέξεων-κλειδιών που συναντώνται στον τίτλο και σε άλλα ζωτικά, ως προς την βελτιστοποίηση, σημεία της ιστοσελίδας αποτελεί την βέλτιστη λύση του κριτηρίου.

Για το παράδειγμα της ιστοσελίδας με τα άνθη-φυτά, η χρήση μιας περιγραφής στην αρχική σελίδα όπως «FlowerShop.GR – Το FlowerShop είναι ένα ηλεκτρονικό κατάστημα που ειδικεύεται στην **πώληση ανθέων και φυτών** και στην αποστολή τους σε όλη την Ελλάδα» είναι επαρκής. Επίσης είναι σημαντικό ο τονισμός των λέξεων «άνθη», «φυτά» στο υπόλοιπο κείμενο της σελίδας χωρίς όμως η μορφοποίηση να γίνει κουραστική γιατί η αισθητική ενός κειμένου παίζει μεγάλο ρόλο στον επισκέπτη.

### **2.2.7 Υπερσυνδέσεις**

Οι υπερσυνδέσεις ή υπερσύνδεσμοι είναι επίσης ένας από τους παράγοντες που επηρεάζουν την βελτιστοποίηση μιας ιστοσελίδας. Ελέγχοντας διάφορες ιστοσελίδες, παρατηρείται ότι δεν δίνεται μεγάλη προσοχή στην σωστή σύνταξη τους και επακόλουθο είναι το χάσιμο πολύτιμων πόντων βελτιστοποίησης. Όπως στις υπόλοιπες ετικέτες που αναφέρθηκαν, οι υπερσυνδέσεις θα πρέπει να είναι σωστά δομημένες.

Το κύριο χαρακτηριστικό των συνδέσεων είναι η διεύθυνση που ανακατευθύνει όταν ο χρήστης την ακολουθεί. Μια μηχανή αναζήτησης διαβάζει την διεύθυνση αυτή και δημιουργεί μια εικόνα για το τι περίπου περιέχει η σελίδα που βρίσκεται από πίσω. Μια ξεκάθαρη διεύθυνση όπως `/flowers/tulips.htm` προϋποθέτει τον χρήστη ότι ακολουθεί μια σελίδα που μιλάει για τουλίπες παρά μια διεύθυνση τύπου `/fakelos/b.html`. Η πρώτη διεύθυνση ονομάζεται φιλική διεύθυνση και έχει αναλυθεί στην αρχή της ενότητας.

Επίσης οι μηχανές αναζήτησης χαρακτηρίζουν ως αδύναμες τις υπερσυνδέσεις που περιέχουν δυναμική διεύθυνση, όπως διευθύνσεις τύπου `seli-`



da.php?id=5. Ο χαρακτηρισμός αδύναμης διεύθυνσης δεν σημαίνει παράβλεψη. Η ευρετηρίαση τους γίνεται κανονικά αλλά είναι φτωχές στην χρησιμότητα τους, για την μηχανή αναζήτησης, αφού δεν χαρακτηρίζουν το περιεχόμενο που αντιπροσωπεύουν. Διευθύνσεις όπως η προηγούμενη χαρακτηρίζονται ως μη φιλικές.

Αξιοσημείωτο είναι ότι μηχανές αναζήτησης δεν μπορούν να διαβάσουν γλώσσες scripting, όπως JavaScript, οπότε η χρήση τους για την δημιουργία δυναμικών συνδέσμων θα πρέπει να αποφεύγεται. Για παράδειγμα, στην JavaScript στο `a href`, θέτεται η διεύθυνση του συνδέσμου με τις εντολές `document.location='διεύθυνση'`. Η μηχανή αναζήτησης όμως δεν μπορεί να ακολουθήσει την σύνδεση, οπότε την παραβλέπει. Γι αυτόν τον σκοπό θα πρέπει ο σύνδεσμος να οδηγεί μέσω του χαρακτηριστικού `href` σε μια υπάρχουσα σελίδα, αντίστοιχη με αυτή που οδηγεί το κομμάτι JavaScript, ενώ στο γεγονός `onclick` να δοθεί η διεύθυνση με κώδικα, ώστε η μηχανή αναζήτησης να αγνοήσει το `onclick` και να διαβάσει μια στατική διεύθυνση.

Ιδιαίτερη βαρύτητα δίνεται στο όνομα των υπερσυνδέσεων. Όταν ένας spider συναντά μια υπερσύνδεση, τότε χρησιμοποιεί το κείμενο της για να σχηματίσει την σχετική φράση-κλειδί για την σύνδεση αυτή. Όταν τα κείμενα είναι περιγραφικά και σχετικά με το περιεχόμενο της σύνδεσης τους, έχουμε καλή βελτιστοποίηση. Για παράδειγμα σε μια ιστοσελίδα που αναφέρεται σε πώληση ακινήτων, η περιγραφή της σύνδεσης ως «Αγορές ακινήτων» δίνει μεγαλύτερη αξία στην υπερσύνδεση από μια απλή σύνδεση με όνομα «ΕταιρίαΑΕ»

Όπως και στις υπόλοιπες ετικέτες της html, γίνεται και εδώ η εκμετάλλευση των λέξεων-κλειδιών σε όλες τις θέσεις του συνδέσμου, από την διεύθυνση, η οποία μπορεί να περιέχει φακέλους και ονόματα αρχείων σχετικά με τις λέξεις-κλειδιά, μέχρι και την χρήση του χαρακτηριστικού `title` που περιέχει την περιγραφή του συνδέσμου στους φυλλομετρητές για AMEA ή αισθητικά εμφανίζει το `title` με την κίνηση του ποντικού πάνω από την σύνδεση.

Οι μηχανές αναζήτησης όταν επισκέπτονται κάποια υπερσύνδεση θέλουν να δέχονται μια σελίδα ως απάντηση και όχι κάποιο σφάλμα. Συχνά παρατηρείται το φαινόμενο του σπασμένου συνδέσμου, μιας υπερσύνδεσης που δεν οδηγεί πουθενά λόγω μη σωστής πληκτρολόγησης της διεύθυνσης ή μετακίνη-

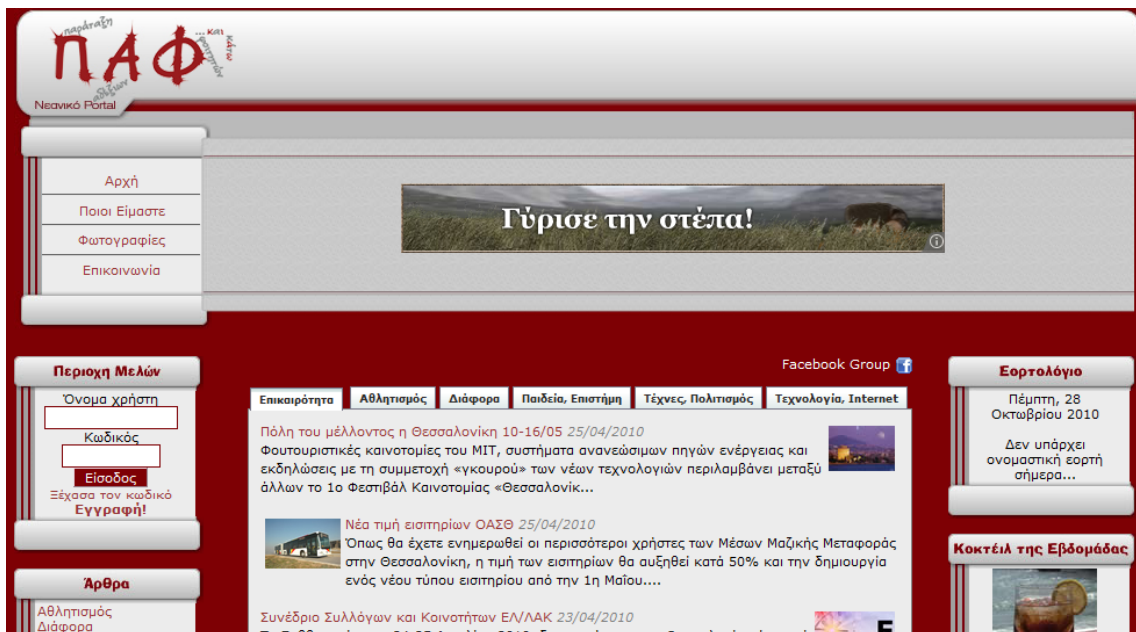
σης της σελίδας που αντιστοιχούσε. Ένας ιδιοκτήτης ιστοσελίδας θα πρέπει να ελέγχει τακτικά τις υπερσυνδέσεις ώστε να μην δεχθεί αρνητική βαθμολογία από τις μηχανές αναζήτησης για σπασμένους σύνδεσμούς.

Τέλος, ορισμένες μηχανές αναζήτησης υποστηρίζουν ένα ειδικό χαρακτηριστικό, το “rel=nofollow”. Το nofollow είναι ένα χαρακτηριστικό το οποίο τοποθετείται στις υπερσυνδέσεις ώστε να τις χαρακτηρίσει ως μη σημαντικές. Όταν μια μηχανή αναζήτησης συναντά τέτοιου είδους σύνδεσμο τον ακολουθεί αλλά εξαιρεί την σύνδεση από τα χαρακτηριστικά που εξετάζονται για την βελτιστοποίηση και δεν επηρεάζει παράγοντες όπως το PR.

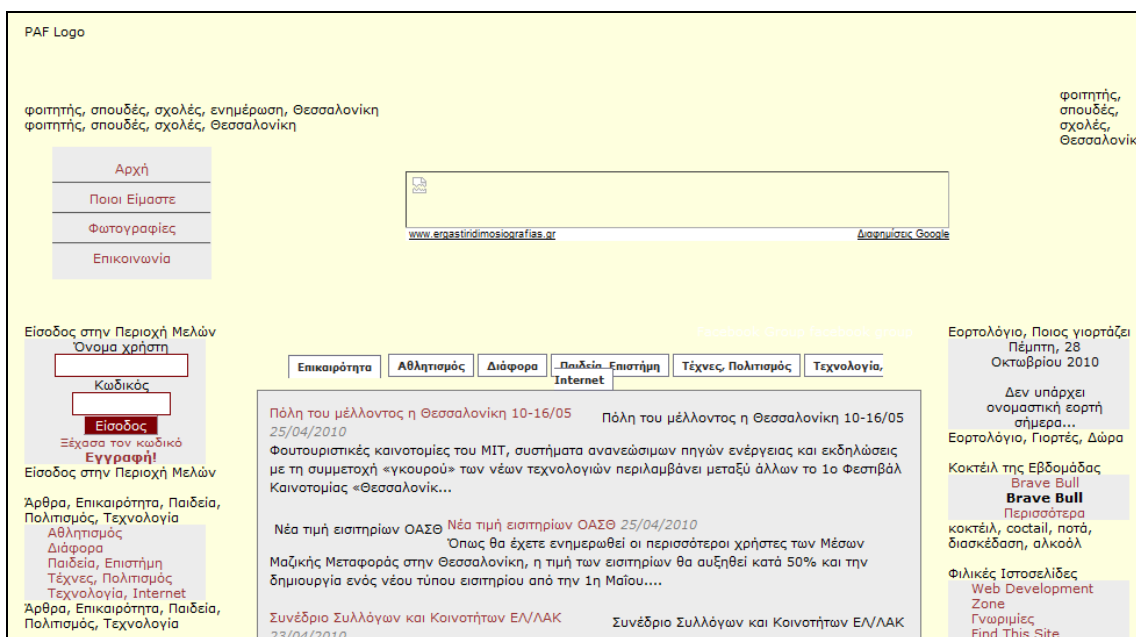
Το χαρακτηριστικό αυτό χρησιμοποιείται σε υπερσυνδέσεις όπως υπερσύνδεση σε μια σελίδα επικοινωνίας ή σελίδα γενικών πληροφοριών και γενικά σε σελίδες με ασήμαντο περιεχόμενο. Επίσης, χρησιμοποιείται σε προγράμματα όπου μπορούν οι επισκέπτες να δημοσιεύουν συνδέσμούς. Ως γνωστόν μια ιστοσελίδα επηρεάζεται από τις συνδέσεις τις ως προς άλλες. Σε περίπτωση που ο χρήστης δημοσιεύει συνδέσμούς δεν είναι εύκολο να ελέγχονται ή να είναι όλοι σχετικοί και ισάξιοι της ιστοσελίδας που τους περιέχει. Η χρήση του χαρακτηριστικού nofollow αποτρέπει την επιρροή του PR και άλλων παραγόντων από τους δημοσιευμένους συνδέσμούς.

## **2.2.8 Alt tags**

Σημαντικό κριτήριο για την σωστή βελτιστοποίηση των εικόνων, όσο αφορά τις αναζητήσεις εικόνων, αλλά και γενικά την σελίδα είναι η χρήση του χαρακτηριστικού alternative. Κάθε εικόνα θα πρέπει να συνοδεύεται από το alt χαρακτηριστικό της. Στο χαρακτηριστικό alt αντιστοιχεί μια περιγραφή της εικόνας που εξυπηρετεί φυλλομετρητές που δεν υποστηρίζουν ή έχουν απενεργοποιημένες τις εικόνες (όπως PDA, κινητά) αλλά και φυλλομετρητές κατάλληλους για άτομα με ειδικές ανάγκες. Εκτός από τα δύο χρηστικές αξίες, οι μηχανές αναζήτησης διαβάζουν το περιεχόμενο του χαρακτηριστικού alt οπότε δίνεται η ευκαιρία στον χρήστη να τις εμπλουτίσει κατάλληλα με λέξεις-κλειδιά. Στην Εικόνα 2.6 εμφανίζεται μια ιστοσελίδα με τις εικόνες της ενώ στην εικόνα Εικόνα 2.7 απεικονίζεται η ίδια σελίδα με απενεργοποιημένες εικόνες. Τα εναλλακτικά κείμενα εμφανίζονται στην θέση των εικόνων .



Εικόνα 2.6: Περιήγηση σε ιστοσελίδα με εικόνες



Εικόνα 2.7: Περιήγηση σε ιστοσελίδα με απενεργοποιημένες εικόνες

Η χρήση μιας περιγραφής, δομημένη με λέξεις-κλειδιά, είναι θεμιτή σε μια εικόνα. Δεν θα πρέπει όμως να βομβαρδίζεται με άσχετες λέξεις-κλειδιά παρά με αυτά που την χαρακτηρίζουν. Ο άσκοπος βομβαρδισμός με λέξεις-κλειδιά μπορεί να ξεγελάσει μια μηχανή αναζήτησης, αλλά αναιρείται η χρήση του χαρακτηριστικού alt για τα άτομα με ειδικές ανάγκες και δυσκολεύει την περιήγηση τους αφού θα διαβάζουν άσχετο, με τις εικόνες, κείμενο. Σύγχρονες μηχανές αναζήτησης ανιχνεύουν τον βομβαρδισμό με λέξεις-κλειδιά και η θεωρούν την

ιστοσελίδα ως keyword spamming. Αν χαρακτηριστεί μια ιστοσελίδα ως spamming σελίδα προσθέεται σε μαύρη λίστα και διαγράφεται μόνιμα από την μηχανή αναζήτησης.

Πρόσφατες μελέτες πάνω στην βελτιστοποίηση ιστοσελίδων δείχνουν ότι η χρήση λέξεων-κλειδιών στις εικόνες δεν υφίσταται ως τεχνική βελτιστοποίησης. Υποστηρίζουν ότι οι μηχανές αναζήτησης αγνοούν τα alt κείμενα όσο αφορά το κομμάτι αξιολόγησης μιας ιστοσελίδας, αλλά το χρησιμοποιούν μόνο για σκοπούς ευρετηρίασης. Άλλοι αναλυτές βελτιστοποίησης υποστηρίζουν ότι η τεχνική ίσως να υπολογίζεται ακόμη και προτείνουν την χρήση μιας σωστής περιγραφής της εικόνας με την χρήση κάποιας λέξης ή φράσης κλειδί.

Για το παράδειγμα η χρήση του ονόματος κάθε άνθους στις εικόνες όπως «Άνθος Παπαρούνα», «Φυτό Λεμονιά» και η περιγραφή των εικόνων γύρω από τις λέξεις άνθη, φυτά, δώρα, αποτελεί μια σωστή χρήση της τεχνικής.

## 2.2.9 Πυκνότητα Λέξεων-Κλειδιών

Η πυκνότητα των λέξεων-κλειδιών είναι ένας ακόμη παράγοντας που πρέπει να ληφθεί σοβαρά υπόψη κατά την βελτιστοποίηση. Ως πυκνότητα λέξεων-κλειδιών ή αλλιώς keyword density ονομάζεται μια ένδειξη του πλήθους των εμφανίσεων μιας συγκεκριμένης λέξης-κλειδιού σε μια σελίδα και εκφράζεται ως ένα ποσοστό επί του συνόλου των λέξεων της σελίδας. Για το κριτήριο της πυκνότητας όλα τα κείμενα της σελίδας χωρίζονται σε μονές λέξεις και καταγράφονται. Όταν μια λέξη υπάρξει για περισσότερες από μια φορά στο κείμενο, προστίθεται στο πλήθος των εμφανίσεων της συγκεκριμένης λέξης και αφού συγκεντρωθούν όλες οι λέξεις, υπολογίζεται ένα ποσοστό εμφάνισης της κάθε λέξης στο συνολικό κείμενο.



Εικόνα 2.8. Σύννεφο βασισμένο στην πυκνότητα λέξεων-κλειδιών της σελίδας webdevzone.gr

Το σύννεφο λέξεων είναι μια αναπαράσταση της πυκνότητας των λέξεων-κλειδιών και δίνει στον χρήστη μια εικόνα για το ποιες λέξεις επαναλαμβάνονται συχνά και διακρίνονται στο κείμενο του. Οι λέξεις με εντονότερο και μεγαλύτερο μέγεθος γραμματοσειράς επαναλαμβάνονται συχνότερα. Υπάρχουν διάφορα εργαλεία στο διαδίκτυο που δημιουργούν αυτά τα σύννεφα.

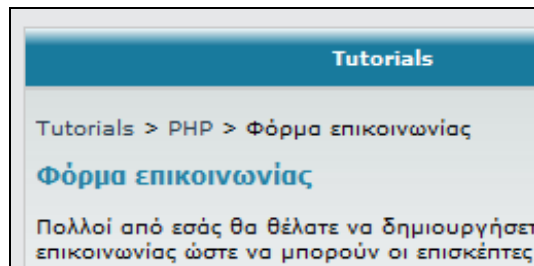
Ιδιαίτερη προσοχή θα πρέπει να δοθεί στην υπερβολική επανάληψη των λέξεων-κλειδιών. Η μεγάλη συχνότητα εμφάνισης των λέξεων-κλειδιών στο κείμενο της σελίδας μπορεί να οδηγήσει σε λάθος συμπεράσματα από την μεριά των μηχανών αναζήτησης. Οι μηχανές αναζήτησης θα θεωρήσουν την ιστοσελίδα μη έγκυρη αλλά αναξιόπιστη και ανεπιθύμητη και θα την αφαιρέσουν εξ ολοκλήρου από τους καταλόγους τους αν εντοπίσουν φαινόμενα υπερβολικής επανάληψης ορισμένων λέξεων ή φράσεων κλειδιών.. Φημολογείται ότι το ασφαλές ποσοστό για την πυκνότητα φθάνει της δέκα ποσοστιαίες μονάδες. Το ασφαλέστερο ποσοστό που προτιμούν οι περισσότεροι κινείται ανάμεσα στο 3% ως το 5%.

### **2.2.10 Πλοήγηση Σελίδας**

Η πλοήγηση σε μια ιστοσελίδα είναι σημαντική γιατί βοηθά τους επισκέπτες να βρουν ο,τι ψάχνουν. Επίσης βοηθά τις μηχανές αναζήτησης να κατανοήσουν ποιο περιεχόμενο θεωρείται σημαντικό από τους δημιουργούς της σελίδας. Αν και τα αποτελέσματα αναζήτησης είναι σε επίπεδο σελίδας, δηλαδή ελέγχουν σελίδα-σελίδα μεμονωμένα, οι μηχανές αναζήτησης θέλουν να γνωρίζουν τον ρόλο που παίζει κάθε σελίδα στον ιστότοπο.

Η πλοήγηση πρέπει να έχει μια φυσική ιεραρχία. Θα πρέπει να πηγαίνει από τα γενικά στοιχεία προς τα ειδικά. Θα πρέπει να αποφευχθεί η δημιουργία πολύπλοκης πλοήγησης, δηλαδή η σύνδεση κάθε σελίδας σε κάθε άλλη, αλλά συνδέσεις που έχουν νόημα. Το βάθος του περιεχομένου θα πρέπει να είναι ρηχό, δηλαδή ο χρήστης θα πρέπει να βρίσκει αυτό που ζητά με δυο, τρία κλικ το πολύ. Οι ιστοσελίδες που είναι σημαντικές θα πρέπει να βρίσκονται σε πιο κεντρικούς φακέλους και όχι κρυμμένες σε υποφακέλους υποφακέλων, γιατί θεωρούνται λιγότερο σημαντικές από τον spider.

Η χρήση κειμένου στην πλοήγηση υπερτερεί έναντι ενός drop-down μενού, JavaScript ή γραφικού μενού, διότι μια μηχανή αναζήτησης διαβάζει πιο εύκολα το κείμενο, παρά κώδικα scripting. Ορισμένες μηχανές αναζήτησης αναγνωρίζουν την γραφική πλοήγηση αλλά δυσκολεύονται να την ερμηνεύσουν.



Εικόνα 2.9: Μονοπάτι

Μια τεχνική είναι η χρήση ενός οριζώντιου μενού για την θέση του κάθε χρήστη, κοινώς ως μονοπάτι (Εικόνα 2.9). Το μονοπάτι θα πρέπει ξεκινά από την αρχική σελίδα και να δείχνει βήμα, βήμα ποιες σελίδες ακολουθήθηκαν ώστε να βρεθεί ο επισκέπτης στην τρέχουσα. Με αυτόν τον τρόπο φαίνεται το βάθος του περιεχομένου αλλά και η σωστή δομή της πλοήγησης καθώς και η μετάβαση από τις γενικές σελίδες στις ειδικές.

### 2.2.11 Καθαρός Κώδικας

Όπως και στην φύση, η καθαριότητα είναι σημαντικός παράγοντας για την σωστή και υγιεινή διαβίωση του ανθρώπου, έτσι και στις ιστοσελίδες ο καθαρός κώδικας αποτελεί αξιόλογο σημείο αλλά και σωστή λειτουργία των διαφόρων υπηρεσιών και εργαλείων.

Όπως είναι γνωστό, μια μηχανή αναζήτησης ευρετηριάζει μια ιστοσελίδα στέλνοντας τους spiders της ώστε να καταγράψουν και να αναλύσουν το περιεχόμενο. Ένας spider όμως είναι ένα πρόγραμμα και δεν μπορεί να δει και να αναλύσει μια ιστοσελίδα όπως το ανθρώπινο μάτι. Μια μηχανή συνεννοείται καλύτερα με μια μηχανή. Ο HTML κώδικας αποτελεί το σημείο επικοινωνίας μεταξύ των spiders και των ιστοσελίδων. Ένας spider παίρνει τον HTML κώδικα μιας ιστοσελίδας, αντλεί πληροφορίες από αυτόν και βγάζει τα συμπεράσματα του για την σελίδα.

Όπως ένα κακογραμμένο κείμενο, χωρίς νεύματα, χωρίς προτάσεις και χωρίς παραγράφους είναι δυσνόητο να το αντιληφθεί εύκολα ο άνθρωπος, έτσι

και οι spiders έχουν απαιτήσεις από μια ιστοσελίδα. Ο HTML κώδικας θα πρέπει να είναι σωστός συντακτικά ώστε να είναι αντιληπτό με την ανάγνωση του, τι στοιχείο είναι το κάθε τι και ως που οριοθετείται. Οι νεότεροι spiders είναι ανεκτικοί σε συντακτικά λάθη και προσπαθούν να μαντέψουν ορισμένα σημεία ή τα αγνοούν.

Όσο αφορά το κομμάτι της βελτιστοποίησης, οι ιστοσελίδες χωρίς συντακτικά λάθη στον κώδικα λαμβάνουν καλύτερη αξιολόγηση από τον αλγόριθμο των μηχανών αναζήτησης γιατί οι spiders μπορούν να κατανοήσουν πλήρως το περιεχόμενο της ιστοσελίδας χωρίς να παιδευτούν για μάντεμα ή αγνόηση μερών της. Θα πρέπει να δοθεί μεγάλη προσοχή στον καθαρό κώδικα, γιατί πολλές από τις τεχνικές που αναφέραμε παραπάνω μπορεί να μην ληφθούν υπόψη από μια μηχανή αναζήτησης γιατί ο spider αδυνατούσε να τις εντοπίσει λόγω του μη σωστά συντεταγμένου κώδικα.

### 2.2.12 Rewrite Engine

Το πρόβλημα των δυναμικών ιστοσελίδων όσο αφορά τις διευθύνσεις αναφέρθηκε προηγουμένως και εστιάζεται στην φιλικότητα των διευθύνσεων που δημιουργούνται. Όπως είναι γνωστό μια δυναμική σελίδα δεν έχει κάποια φυσική διεύθυνση αλλά δυναμική, λόγω της χρήσης μεταβλητών στην διεύθυνση της και η αντιστοίχηση τους σε μεταβλητές προγραμματισμού με μεθόδους ανάθεσης τιμών όπως η GET.

Λύση στο πρόβλημα των μη φιλικών διευθύνσεων είναι η χρήση της Rewrite Engine. Ο μηχανισμός αυτός διαβάζει την διεύθυνση όπως την εισήγαγε ο χρήστης και, με τα κριτήρια που δηλώθηκαν στις ρυθμίσεις του web server, μετασχηματίζει την διεύθυνση σε πραγματική διεύθυνση για τον server. Για παράδειγμα, ο χρήστης πληκτρολογεί *articles/read/what-is-seo* και ο μηχανισμός το μετασχηματίζει σε */articles.php?read=what-is-seo*. Η Rewrite Engine υποστηρίζεται από τα εξής λογισμικά, στις βασικές εκδόσεις, αλλά και σαν επεκτάσεις από frameworks.

- Apache HTTP Server
- Microsoft IIS
- Ruby on Rails

- Java Server Faces
- Django
- Java Stripes Framework
- Perl frameworks

### **Apache MOD Rewrite**

Η λύση στον Apache web server, έναν από τους πιο δημοφιλή web servers στο διαδίκτυο, είναι η χρήση του module rewrite, γνωστό ως mod\_rewrite. Οι κανόνες μετασχηματισμού γράφονται με κανονικές εκφράσεις (regular expressions) και μπορούν να συνυπάρχουν με άλλους κανόνες.

Για την δημιουργία ενός κανόνα πρέπει να τροποποιηθεί ένα αρχείο .htaccess που βρίσκεται στον κεντρικό κατάλογο, ρίζα, του web server. Αν το αρχείο δεν υπάρχει μπορεί να δημιουργηθεί από τον χρήστη.

Ένα .htaccess αρχείο είναι ένα αρχείο κειμένου που υποστηρίζεται από τον Apache και περιέχει διάφορες παραμετροποιήσεις και υπερβάσεις των κανόνων και ρυθμίσεων του. Χρησιμοποιείται για να δοθούν προσωπικές ρυθμίσεις σε κάθε φάκελο ή ιστοτόπο ενός web server διότι οι βασικές ρυθμίσεις δεν είναι προσπελάσιμες σε όλους τους χρήστες για λόγους ασφαλείας. Το .htaccess περιέχει μόνο τους κανόνες που αλλάζουν ή προστίθενται για τον ιστοτόπο ή τον φάκελο που το περιέχει χωρίς να επηρεάζονται άλλοι ιστοτόποι.

Για την δημιουργία ενός αρχείου .htaccess χρησιμοποιείται κυρίως το Notepad ή άλλο πρόγραμμα που αποθηκεύει καθαρό κείμενο, χωρίς μορφοποιήσεις. Ο παρακάτω κώδικας ενεργοποιεί την μηχανή επανεγγραφής

```
RewriteEngine on
Options +FollowSymlinks
RewriteBase /
RewriteRule
```

Έπειτα δηλώνονται οι κανόνες μετασχηματισμού. Για παράδειγμα, μια ιστοσελίδα περιέχει την μη φιλική διεύθυνση `index.php?article=5&page=2` και πρέπει να μετασχηματιστεί σε μια φιλική διεύθυνση, όπως ή `index/article/5/page/2`. Ο κανόνας που μετατρέπει την παραπάνω διεύθυνση στην φιλική της μορφή είναι ο:



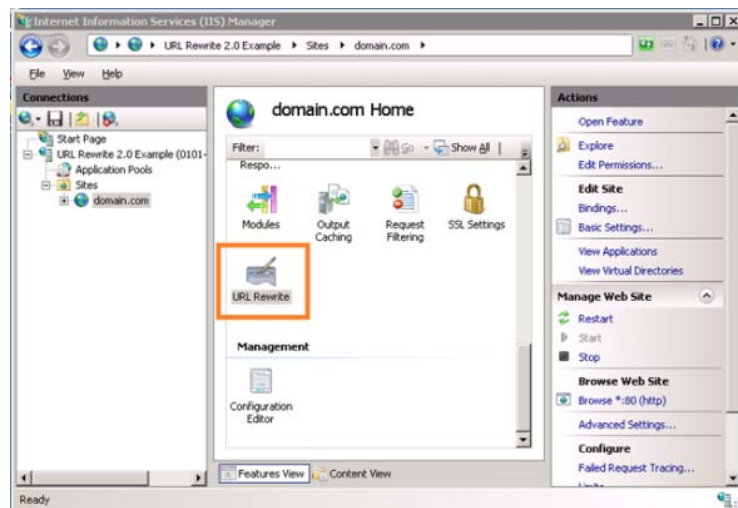
$\wedge(.*)index/article/(.*)/page/(.*)\$ \$1index.php?article=$2&page=$3$

Ο τρόπος μετασχηματισμού είναι ο εξής: Ο χρήστης πληκτρολογεί *www.selida.gr/index/article/2/page/5*. Ο web server μόλις διαβάσει την διεύθυνση, ελέγχει αν η μηχανή επανεγγραφής διευθύνσεων είναι ενεργή. Σε περίπτωση που είναι, ελέγχει αν η διεύθυνση αντιστοιχεί σε κάποιο κανόνα. Εφόσον αντιστοιχεί αποθηκεύει το 2 και το 5 στις μεταβλητές κανονικών εκφράσεων και μετασχηματίζει την διεύθυνση σε αυτή που επιβάλλει ο κανόνας. Κάθε μεταβλητή που χρησιμοποιήθηκε καλείται με \$1,\$2 και δημιουργείται ο πραγματική διεύθυνση ώστε να κληθεί το κατάλληλο αρχείο με τα κατάλληλα ερωτήματα.

Στην ουσία χρησιμοποιούνται οι κανονικές εκφράσεις έτσι ώστε να διασπαστεί η διεύθυνση σε διάφορα μέρη και να ανατεθούν τα μεταβλητά μέρη σε μεταβλητές κανονικών εκφράσεων και καλούνται στην νέα φιλική διεύθυνση με κατάλληλη τοποθέτηση.

### **Microsoft IIS**

Παρόμοια είναι η χρήση της Rewrite Engine στον web server της Microsoft. Σε αντίθεση όμως με τον Apache οι διαδικασίες εισαγωγής των κανόνων στον IIS είναι πιο απλές και αυτοματοποιημένες .



Εικόνα 2.10: Πίνακας ρυθμίσεων IIS

Ο χρήστης θα πρέπει να πλοηγηθεί στο **IIS Manager > Επιθυμητό Website > URL Rewrite** και μετά να επιλέξει τον τρόπο θα δημιουργήσει κανόνα (Εικόνα 2.10). Αν θέλει μπορεί να δημιουργήσει κανόνα από το μηδέν χρησιμοποιώντας κανονικές εκφράσεις, όπως στον Apache, με την επιλογή **Blank**

**Rule.** Αν προτιμά τις αυτοματοποιημένες διαδικασίες, ένα από τα μεγάλα πλεονεκτήματα όλων των λογισμικών της Microsoft, επιλέγει το **User Friendly URL**. Σε αυτήν την επιλογή ο χρήστης εισάγει την πραγματική διεύθυνση και το πρόγραμμα του IIS προτείνει τον χρήστη κάποιες πιθανές φιλικές διευθύνσεις. Ο χρήστης μπορεί να επιλέξει όποια από τις προτεινόμενες καθώς και να τις τροποποιήσει.

## 2.3 Κακόβουλη Βελτιστοποίηση

Όπως αναφέρθηκε και στην εισαγωγή υπάρχει και η μαύρη πλευρά της βελτιστοποίησης. Ακολουθούνε διάφορες τεχνικές που δεν πρέπει να υιοθετούνται γιατί η ιστοσελίδα κινδυνεύει με αποκλεισμό από τους καταλόγους μηχανών αναζήτησης. Υπάρχουν, επίσης, τεχνικές που μπορούν να βλάψουν ιστοσελίδες ανταγωνιστών. Η αναφορά τους έχει *εκπαιδευτικό* και μόνο σκοπό, για την προστασία των ιστοσελίδων καθώς και παραδείγματα προς αποφυγή.

### 2.3.1 Redirect Hijacking

Το page hijacking είναι μια τεχνική που επηρεάζει τον τρόπο που οι μηχανές αναζήτησης ερμηνεύουν συγκεκριμένες εντολές που ένας web server στέλνει στον χρήστη. Στην ουσία, επιτρέπει έναν ιστότοπο να αντικαταστήσει σελίδες από έναν άλλον ιστότοπο στα αποτελέσματα αναζήτησης, χωρίς την άδεια και την γνώση του δεύτερου.

Όταν ένας επισκέπτης αναζητά κάποιον όρο, η μηχανή αναζήτησης επιστρέφει κάποιες ιστοσελίδες. Ένας hijacker μπορεί να αντικαταστήσει τις σελίδες που εμφανίζονται στα αποτελέσματα με ιστοσελίδες που αυτός/η ελέγχει. Οι νέες σελίδες που εισάγει ο hijacker είναι εικονικές σελίδες, κυρίως server side scripts, και δεν υπάρχουν στην πραγματικότητα. Ο επισκέπτης οδηγείται απευθείας από τα αποτελέσματα αναζήτησης στο script του hijacker. Μια σελίδα που δέχθηκε page hijacking εμφανίζεται με σωστή περιγραφή στα αποτελέσματα αναζήτησης αλλά διαφορετική διεύθυνση.

Όταν κάποιος θέλει να κάνει redirect hijack ανακατευθύνει τον χρήστη σε άλλες σελίδες που αυτός επιλέγει. Αν το hijacking είναι κρυφό και δεν μπορεί να εντοπιστεί εύκολα από τα spiders των μηχανών αναζήτησης, μπορεί να διαρκέσει μεγάλο χρονικό διάστημα.

Ένα πιθανό σενάριο redirect hijacking είναι η εμφάνιση ιστοσελίδων πορνογραφικού περιεχομένου ως σελίδες ειδήσεων του CNN, δημιουργία ψευδών πρωτοσέλιδων ηλεκτρονικών καταστημάτων, τραπεζών κτλ

Το page hijacking γίνεται ως εξής. Ο hijacker προσπαθεί να εισάγει κάποιο 302 redirect ή meta refresh σε μια ιστοσελίδα. Αφού το επιτύχει προσπαθεί να κοροϊδέψει τον spider και να γίνει το script του επίσημη διεύθυνση μιας ιστοσελίδας ενός άλλου webmaster. Πλέον, οι χρήστες θα βλέπουν τον σωστό τίτλο και περιγραφή της ιστοσελίδας που δέχθηκε hijacking αλλά λάθος διεύθυνση.

Ανάλογα με την επιτυχία των hijackers μια ιστοσελίδα θα πληγεί από τους επισκέπτες που δέχεται από τις μηχανές αναζήτησης και θα εξαφανιστεί σιγά σιγά γιατί όλες οι σελίδες της είναι μολυσμένες και δεν θα εμφανίζονται σε σχετικές αναζητήσεις. Δυστυχώς οι ιστοσελίδες που δέχονται hijacking δεν μπορούν να αμυνθούν εφόσον ο spider έχει ευρετηριάσει το script ως διεύθυνση της ιστοσελίδας. Ο μόνος τρόπος επίλυσης είναι το έλεος που μπορεί να δείξει ο hijacker και να μετατρέψει την διεύθυνση του script σε μια 404 σελίδα και να διαγραφεί από το ευρετήριο.

Πολλές μηχανές αναζήτησης είναι ευάλωτες στην επίθεση με page/redirect hijack συμπεριλαμβανομένων της Google και της Bing. Η Yahoo! είναι η μόνη μηχανή αναζήτησης που κατάφερε να κλείσει την τρύπα του hijacking.

Πολλές φορές μια ιστοσελίδα μπορεί να γίνει hijacker κατά λάθος. Γι αυτό υπάρχει μια σειρά από ενέργειες που δεν πρέπει να κάνει.

Σε περίπτωση που κάποια script χρησιμοποιούν 302 redirects πρέπει να μετατρέπονται σε 301 redirect. Ένα 302 redirect σημαίνει προσωρινή αλλαγή της διεύθυνσης. Δίνει την οδηγία στην μηχανή αναζήτησης να διαβάσει και να χρησιμοποιήσει το περιεχόμενο της νέας σελίδας αλλά να ελέγχει τακτικά την αρχική διεύθυνση γιατί θα πρόκειται να αλλάξει πάλι. Το 301 redirect είναι μόνιμη αλλαγή διεύθυνσης και δεν μπορεί να επηρεαστεί από hijacking

Σε περίπτωση που δεν είναι δυνατή η αλλαγή του τύπου redirect, πρέπει τα redirect scripts να απενεργοποιηθούν στο "robots.txt". Το robots.txt είναι ένα αρχείο κειμένου που περιέχει οδηγίες για τους spiders των μηχανών αναζήτη-

σης. Αν ο ιδιοκτήτης μιας ιστοσελίδας αντιληφθεί ότι έχει κάνει hijack σε μια άλλη ιστοσελίδα στο Google, πρέπει να αλλάξει το script ώστε να επιστρέφει σφάλμα 404 και να ζητήσει να αφαιρεθεί το script από το ευρετήριο της μηχανής αναζήτησης.

### 2.3.2 Page Swapping (Cloaking)

Η ανταλλαγή σελίδων, ή αλλιώς page swapping ή cloaking, αποτελεί μέρος του μαύρου SEO που αναλύεται για εγκυκλοπαιδικούς σκοπούς και δεν παροτρύνεται η χρήση του. Είναι μια τεχνική στην οποία τοποθετείται μια πολύ καλά βελτιστοποιημένη σελίδα ώστε η εκάστοτε ιστοσελίδα να πάρει υψηλές βαθμολογίες και μετά η αντικατάσταση του από χαμηλότερης ποιότητας σελίδα όσο αφορά την βελτιστοποίηση αλλά καλύτερου οπτικού περιεχομένου για τον χρήστη.

Πιο αναλυτικά, με την έναρξη ενός καινούριου ιστοτόπου τοποθετείται σαν αρχική του σελίδα μια σελίδα απλή, βελτιστοποιημένη στο μέγιστο με όλους τους τρόπους που αναφέρθηκαν παραπάνω και ίσως μερικές μαύρες ή γκρίζες SEO τεχνικές. Ο spider που θα την επισκεφτεί και θα την αναλύσει, θα δει ότι πληροί πολλά από τα κριτήρια βελτιστοποίησης οπότε θα την αξιολογηθεί θετικά και ποιοτικά. Αφού γίνει η καταγραφή και η αξιολόγηση, ο διαχειριστής του ιστοτόπου αντικαθιστά την σελίδα με αυτήν που προοριζόταν εξαρχής. Ως αποτέλεσμα, ο ιστότοπος να έχει αξιολογηθεί για ποιοτικός από την προηγούμενη σελίδα, ενώ η τρέχουσα είναι κατώτερη και εντελώς διαφορετική από αυτή που διάβασε ο spider. Υπάρχουν δυο τρόποι cloaking, το *User Agent* και το *IP Agent Cloaking*

#### ***User Agent***

Όταν ένας επισκέπτης, μια μηχανή αναζήτησης ή ένας χρήστης, εισέρχεται σε έναν ιστότοπο, το script για cloaking διαβάζει το όνομα του User Agent και το συγκρίνει με μια λίστα των πιο γνωστών ονομάτων User Agents των μηχανών αναζήτησης. Συνήθως, όταν ένας spider κουβαλά μαζί του και το όνομα της μηχανής αναζήτησης που το στέλνει. Αν βρεθεί μέσα στην λίστα τότε το script του σερβίρει την βελτιστοποιημένη, αλλά φτωχή οπτικά, σελίδα. Αν όμως ο spider δεν φέρει όνομα, το script τον θεωρεί άνθρωπο και του σερβίρει την σελίδα που είναι σχεδιασμένη για καλύτερη οπτική εικόνα.

### ***IP Agent***

Αυτή η μέθοδος είναι πιο πολύπλοκη σε σχέση με την προηγούμενη, όσο αφορά τον εντοπισμό του επισκέπτη, γιατί περιλαμβάνει τις διευθύνσεις IP των spiders. Με την ίδια ακριβώς μέθοδο, όταν κάποιος επισκέπτεται τον ιστότοπο, ελέγχεται η διεύθυνση IP του. Αν ανήκει στην λίστα των διευθύνσεων IP των spider, τότε σερβίρεται η βελτιστοποιημένη σελίδα. Αλλιώς σερβίρεται η κανονική σελίδα. Αυτή η μέθοδος είναι πιο αποτελεσματική γιατί δεν εξαρτάται από το αν ο spider περιλαμβάνει το όνομα της μηχανής αναζήτησης που ανήκει.

### ***Εντοπισμός cloaking***

*Μέθοδος User Agent:* Η μηχανή αναζήτησης στέλνει διάφορους spiders σε μια ιστοσελίδα. Μερικοί περιέχουν το όνομα της μηχανής αναζήτησης και άλλοι όχι. Αν οι spiders εντοπίσουν διαφορετικές σελίδες - αυτή που σερβίρεται στον spider με το όνομα και αυτή στον άλλον που συγχέεται με άνθρωπο – θεωρούν ότι υπάρχει cloaking.

*Μέθοδος IP Agent:* Όταν μια μηχανή αναζήτησης στέλνει έναν spider σε μια ιστοσελίδα από μια διαφορετική διεύθυνση IP από αυτές που συνήθως έστειλε παλιά, είναι πολύ πιθανόν η νέα διεύθυνση να μην υπάρχει στην βάση των IPs. Με την ίδια διαδικασία όπως η παραπάνω, αν εντοπιστούν διαφορετικές σελίδες ανάλογα με την IP που έχει ο spider, θεωρείται cloaking.

*Ανθρώπινη παρέμβαση:* Επειδή αρκετές φορές τα μηχανήματα κάνουν και λάθη απαιτείται η ανθρώπινη παρέμβαση. Κάποιος εργαζόμενος από την μηχανή αναζήτησης επισκέπτεται τον ιστότοπο που πιθανόν να έχει cloaking και ελέγχει αν η σελίδα που προβάλλεται στην οθόνη του είναι τελείως διαφορετική από αυτήν που βλέπει ο spider. Αν εντοπιστεί απάτη, γίνονται οι κατάλληλες ενέργειες για απομάκρυνση της σελίδα από την μηχανή αναζήτησης.

### ***Reverse cloaking***

Πολλές φορές ο DNS Server δείχνει σε δύο ή περισσότερες ιστοσελίδες που φιλοξενούνται στην ίδια διεύθυνση IP. Αυτό ονομάζεται αντίστροφο IP cloaking ή reverse IP cloaking. Ο λόγος που κάποιος θα φιλοξενούσε πολλές ιστοσελίδες σε μια διεύθυνση IP είναι για να κερδίσει περισσότερα back links από όλες τις ιστοσελίδες. Όσο μεγαλύτερο μέρος των σελίδων των ιστοτόπων είναι αλληλένδετοι, τόσο περισσότερο επηρεάζεται αρνητικά την βελτιστοποίηση.

Για να μην υπάρξει η παραπάνω περίπτωση και ποινή των ιστοτόπων θα πρέπει κάθε ιστότοπος να φιλοξενείται σε διαφορετικές διευθύνσεις IP και είναι αποδεκτό να είναι συνδεδεμένες ορισμένες σελίδες των διαφόρων ιστοτόπων μεταξύ τους.

### 2.3.3 Doorway & Σελίδες Πληροφοριών

Η σελίδα doorway, εκ των αγγλικών λέξεων θύρα και δρόμος, ή σελίδα πληροφοριών - στην βιβλιογραφία συναντάται και ως εισαγωγική σελίδα – είναι μια αρχική σελίδα ενός ιστότοπου η οποία καλωσορίζει τον επισκέπτη στον ιστότοπο και δίνει λίγες πληροφορίες γι αυτόν. Οι ενέργειες του χρήστη είναι να πατήσει κάποιο σύνδεσμο ώστε να μπει στην κυρίως σελίδα.

Από σκοπιά βελτιστοποίησης, είναι ένας αθέμιτος τρόπος γιατί η σελίδα αυτή είναι πολύ βελτιστοποιημένη και αποσκοπεί την προσέλκυση των μηχανών αναζήτησης και όχι στον υποτιθέμενο σκοπό του καλωσορίσματος του χρήστη και την πληροφόρηση του για την σελίδα.

Η μορφή των σελίδων αυτών μοιάζει. Στις περισσότερες doorways υπάρχει έμφαση στις βασικές λέξεις-κλειδιά της ιστοσελίδας. Εντοπίζεται κυρίως ένας βελτιστοποιημένος σύνδεσμος εισόδου στην κεντρική σελίδα και από κάτω μια παράγραφο με μικρή περιγραφή που περιέχει λέξεις-κλειδιά. Πολλές φορές υιοθετούνται και άλλες τεχνικές όπως κρυφό κείμενο, keyword spamming και άλλες μαύρες τεχνικές βελτιστοποίησης.

### 2.3.4 Κρυφό Κείμενο

Μια τεχνική που χρησιμοποιείται μέχρι σήμερα είναι η χρήση των κρυφών κειμένων. Γίνεται μεγάλη προσπάθεια να αντιμετωπιστεί από τις μηχανές αναζήτησης αλλά υπάρχουν ποικίλοι τρόποι να κρυφτεί κάποιο κείμενο και να μην εντοπιστεί από τον spider. Το κρυφό κείμενο περιέχει κυρίως λέξεις-κλειδιά ή γενικά μια παράγραφο πλήρως βελτιστοποιημένη ώστε να δώσει αύξηση στον γενικό δείκτη βελτιστοποίησης.

Από τους τρόπους απόκρυψης κειμένου είναι η χρήση της εντολής *display: none* στο CSS μέσα στην HTML. Η εντολή αυτή δεν εμφανίζει το κείμενο

αλλά ότι δεν φαίνεται σε μια σελίδα, δεν σημαίνει ότι δεν υπάρχει και στον κώδικα της. Οι spiders διαβάζουν τον κώδικα, οπότε επηρεάζονται από αυτό.

Ένας άλλος τρόπος είναι η χρήση κοινού φόντου και χρώμα γραμματοσειράς. Σε πολλές ιστοσελίδες που το χρησιμοποιούν, παρατηρείται κυρίως στο κάτω μέρος ένα κενό. Στο κενό εκείνο είναι γραμμένη η παράγραφος ή οι λέξεις-κλειδιά με το ίδιο χρώμα που είναι και το φόντο. Αν μαρκαριστεί με το ποντίκι, το κείμενο γίνεται διακριτό.

Τέλος η χρήση των divs & των εικόνων σε πρώτο πλάνο αποτελούν και αυτοί τρόπους κρυφού κειμένου. Το βελτιστοποιημένο κείμενο μπορεί να εισαχθεί σε ένα div που αναφέρεται πρώτο στον κώδικα και με την χρήση των CSS και της εντολής position: absolute και αρνητική θέση (πχ left -10px) το div μετακινείται εκτός της οθόνης, πράγμα που το κάνει μη αντιληπτό. Παρομοίως οι εικόνες επικαλύπτουν το κείμενο και δεν φαίνεται.

### **2.3.5 Άλλες "Μαύρες" Τεχνικές**

Ακολουθούν μερικές πιο απλές τεχνικές που χρησιμοποιούνται για να βλάψουν μια ιστοσελίδα ή ίδια της να καταχραστεί τους κανόνες της βελτιστοποίησης.

#### ***Στέλνοντας κακό Κάρμα***

Μερικοί webmasters χρησιμοποιούν ανακατευθύνσεις (redirects) όπως με το meta refresh έτσι ώστε να κάνουν ανακατεύθυνση από κάποιο banned domain σε σελίδα των ανταγωνιστών. Ο κάτοχος ενός απαγορευμένου domain (λόγω παράβασης) μπορεί να στέλνει επισκέπτες σε ιστοσελίδες τρίτων με σκοπό να μπερδέψει τους spiders και να δείξει ότι η σελίδα που αποκλείστηκε έχει άμεση σχέση με αυτήν που στέλνει επισκέπτες ώστε να επηρεαστεί αρνητικά η υγιής σελίδα.

#### ***Spamming κειμένου σε σύνδεσμο***

Ο τρόπος αυτός υλοποιείται σχετικά εύκολα και μπορεί να επηρεάσει αρνητικά μια ιστοσελίδα από δημοφιλείς όρους αναζήτησης. Για να γίνει αυτό, ο κακόβουλος webmaster δημιουργεί σε διάφορες σελίδες συνδέσμους με όμοιο κείμενο ανάμεσα στα anchor tags και σύνδεση στην σελίδα του ανταγωνιστή. Αυτό οδηγεί τις μηχανές αναζήτησης να θεωρήσουν τον σύνδεσμο ως καταχραστή και να

επιβάλουν κυρώσεις στην σελίδα που παραπέμπει, χωρίς να γνωρίζει ο ιδιοκτήτης της της τίποτα.

### ***Blogspam***

Υπάρχουν προγράμματα που αναλαμβάνουν να επισκέπτονται διάφορα blogs και να απαντούν στις αναρτήσεις τους άσκοπα με την εισαγωγή ενός συνδέσμου που έχει δοθεί στο πρόγραμμα. Κάποιος κακόβουλος μπορεί να στείλει την διεύθυνση του ανταγωνιστή του σε διάφορα blogs με αυτόν τον τρόπο, αμαυρώνοντας την φήμη της ιστοσελίδας από χρήστες αλλά και μηχανές αναζήτησης, ως spam.

### ***Link Farms, Forums, Βιβλία Επισκεπτών & Spam***

Όπως έχει αναφερθεί και στην εισαγωγή, οι φάρμες συνδέσμων και γενικά το spamming σε forums, βιβλία επισκεπτών έχουν πλέον καταργηθεί και δεν βοηθούν καθόλου στην βελτιστοποίηση της ιστοσελίδας. Μερικοί χρήστες χρησιμοποιούν φθηνά bots για να κάνουν spam σε forums και βιβλία επισκεπτών υποτιθέμενοι ότι απαντούν σε ερωτήματα ή γράφουν καλά σχόλια για την σελίδα του επισκέπτη.

Από την έναρξη των φορμών μέχρι σήμερα πολλοί κατάλογοι ιστοσελίδων δεν έχουν ποιοτικά στάνταρντ και προσθέτουν όλους τους συνδέσμους που καταχωρούνται από τους χρήστες. Δεν υπάρχει οργανωμένο περιεχόμενο ούτε ταξινόμηση σε κατηγορίες και αποδοχή μόνο των σωστών και αξιόλογων συνδέσμων. Μερικοί ιδιοκτήτες τέτοιων καταλόγων τους δημιούργησαν με σκοπό το κέρδος και την πώληση καταχωρήσεων όπως παλιά.

Με το σκεπτικό των παραπάνω τεχνικών κακού κάρματος, έτσι και εδώ η αποστολή της ανταγωνιστικής ιστοσελίδας σε τέτοιους ιστοτόπους χαλάει το όνομα της και την ποιότητα του περιεχομένου της.



# 3 Λύσεις Βελτιστοποίησης

Όπως ειπώθηκε και στην εισαγωγή, η βελτιστοποίηση ιστοσελίδων αποτελεί σημαντικό παράγοντα για κάθε ιστοσελίδα αν θέλει να πετύχει και να ανταγωνιστεί άλλες σελίδες. Ωστόσο, μια ιστοσελίδα μπορεί να επιτύχει μόνο από το περιεχόμενο της και τη μοναδικότητα της όπως το [www.zoo.gr](http://www.zoo.gr). Η βελτιστοποίηση την βοηθά ώστε να γίνει πιο ευδιάκριτη στα αποτελέσματα των μηχανών αναζήτησης και να αποκτήσει κοινό που δεν την γνώριζε.

Στην αγορά υπάρχουν διάφορες λύσεις βελτιστοποίησης. Όταν ένας χρήστης είναι άπειρος σε θέματα βελτιστοποίησης, υπάρχουν εταιρίες όπου ειδικεύονται στην βελτιστοποίηση ιστοσελίδων και αναλαμβάνουν την βελτιστοποίηση αλλά και την παρακολούθηση της προόδου μιας ιστοσελίδας. Αν όμως ο χρήστης θέλει να ασχοληθεί μόνος του για οικονομικούς ή όχι λόγους, υπάρχουν πληθώρα εργαλείων και οδηγών.

## 3.1 Βελτιστοποίηση κατά τη Σχεδίαση

Η επιλογή της βελτιστοποίησης κατά την σχεδίαση αποτελεί την βέλτιστη λύση που κυκλοφορεί στο εμπόριο και αποδίδει περισσότερο από όλες. Η δημιουργία της ιστοσελίδας εξ' αρχής στα βήματα της βελτιστοποίησης βοηθά στην επίτευξη όλων των τεχνικών που προ-αναφέρθηκαν λόγω της ελευθερίας στη συγγραφή κώδικα. Η λύση επιλογή εφαρμόζεται στην δημιουργία νέων ιστοσελίδων αλλά και όταν εφαρμόζεται ολική ανακατασκευή του περιεχομένου της ιστοσελίδας.

Στο ξεκίνημα μιας ιστοσελίδας επιλέγεται το επιθυμητό όνομα domain. Λόγω της μεγάλης συμβολής του ονόματος στην βελτιστοποίηση, πρέπει να επιλεγεί προσεχτικά. Εκτός του ονόματος όμως, θα πρέπει να δοθεί προσοχή και στην κατάληξη του domain για την γεωγραφική βελτιστοποίηση. Χρησιμοποιούνται κυρίως βασικά λέξεις-κλειδιά για την δημιουργία του ονόματος.

Μετά το στήσιμο της βάσης, που είναι το domain, ξεκινάει το χτίσιμο της ιστοσελίδας. Αφού συγκεντρωθούν οι βασικές λέξεις-κλειδιά, ακολουθούνται όλες οι τεχνικές της βελτιστοποίησης που αναφέρθηκαν, όπως τίτλος, meta tags,

alt tags σύμφωνα με την καλαισθησία ως προς τον χρήστη αλλά και την δύναμη που κρύβουν για την σωστή βελτιστοποίηση. Οι τεχνικές που εφαρμόζονται είναι γνωστές ως On Page και Off Page τεχνικές βελτιστοποίησης. Τέλος παρακολουθείτε η πρόοδος της ιστοσελίδας από προγράμματα στατιστικών και γίνονται τυχόν διορθώσεις.

Τα αποτελέσματα της λύσης είναι μέγιστα διότι εφαρμόζονται όλοι οι κανόνες και δεν υπάρχει κάποιος περιορισμός. Σε περίπτωση μίσθωσης για την υπηρεσία βελτιστοποίησης, το χρηματικό πόσο είναι αρκετά υψηλό διότι χρεώνεται η δημιουργία μιας νέας ιστοσελίδας και η ταυτόχρονη μελέτη της.

## 3.2 Βελτιστοποίηση εκ των Υστέρων

Η δεύτερη λύση είναι η βελτιστοποίηση υπαρχών ιστοσελίδων ώστε να συμφωνούν με τους κανόνες σωστής βελτιστοποίησης. Εφαρμόζεται κυρίως σε σελίδες που προ-υπάρχουν στο διαδίκτυο και εφαρμόζεται εύκολα σε στατικές ιστοσελίδες ενώ η χρήση της σε δυναμικές ιστοσελίδες μπορεί να σημαίνει επανακωδικοποίηση του συστήματος και αναδιάταξη του περιεχομένου της.

Σε αντίθεση με την βελτιστοποίηση κατά την σχεδίαση, η λύση αυτή θέτει πολλούς περιορισμούς ως προς τις τεχνικές οι οποίες μπορούν να εφαρμοστούν. Για παράδειγμα, το domain name μπορεί να μην είναι κατάλληλο ή να μην περιέχει λέξεις-κλειδιά. Παρόλο που η αλλαγή ενός domain name δεν είναι δύσκολη, όσο αφορά τα διαδικαστικά, αποφεύγεται διότι όταν μια ιστοσελίδα βρίσκεται στο διαδίκτυο για αρκετό διάστημα είναι πολύ πιθανόν να υπάρχουν άλλες ιστοσελίδες που έχουν κάποια υπερσύνδεση ως προς αυτήν αλλά και οι επισκέπτες έχουν απομνημονεύσει την διεύθυνση της. Τέλος το domain name μπορεί να αποτελεί την ονομασία της ιστοσελίδας και η αλλαγή ονόματος αλλά και διεύθυνσης μπορεί να έχει αρνητικές επιπτώσεις στην επισκεψιμότητα της.

Εκτός των δυσκολιών που συναντώνται στο domain, είναι εξίσου δύσκολο η επανεξέταση και η επανεγγραφή όλου του υλικού σύμφωνα με τους τρόπους που υποδεικνύει η βελτιστοποίηση. Για παράδειγμα, η ενεργοποίηση και η χρήση της Rewrite Engine σημαίνει χειροκίνητη αλλαγή όλων των συνδέσμων που υπάρχουν σε μια ιστοσελίδα. Επίσης υπερσυνδέσεις άλλων ιστοσελίδων σε συγκεκριμένη σελίδα δική σας θα οδηγεί σε μια σελίδα που δεν υπάρχει. Ε-

πίσης, το σύστημα μπορεί να μην έχει προδιαγραφές για την σωστή εφαρμογή της βελτιστοποίησης όπως αυτές που αναφέρθηκαν στο κεφάλαιο 2.2.7

Συνοψίζοντας, τα αποτελέσματα της λύσης αυτής είναι παρόμοια με την πρώτη λύση. Δυστυχώς όμως λόγω των περιορισμών δεν έχουμε την 100% εφαρμογή των τεχνικών που αναφέρθηκαν αλλά που μπορεί να προκύψουν στο μέλλον. Αυτό όμως δεν σημαίνει ότι η βελτιστοποίηση δεν θα έχει αποτελέσματα, απλά θα διαρκέσει περισσότερο καιρό και ίσως κοστίζει παραπάνω λόγω των εργασιών που θα ζητηθούν.

### **3.3 Ατομική Εργασία ή Μίσθωση Επαγγελματία**

Όπως στις περισσότερες πτυχές της σχεδίασης και ανάπτυξης ιστοσελίδων, έτσι και στην βελτιστοποίηση πρωτοστατεί η διαμάχη ανάμεσα σε αυτούς που πιστεύουν ότι η μάθηση και η πρακτική της βελτιστοποίησης θα πρέπει να γίνει από τους ίδιους για τους ίδιους ατομικά και άλλοι που πιστεύουν ότι πρέπει να ανατεθεί σε επαγγελματία. Ακλουθούν τα πλεονεκτήματα και των δυο πλευρών.

#### **Επαγγελματίας**

- Ένας επαγγελματίας με αρκετά χρόνια πείρας στην βελτιστοποίηση ιστοσελίδων θα μπορούσε να ενημερώσει καλύτερα τον ενδιαφερόμενο για το τι θα περιμένει από τις μηχανές αναζήτησης μετά την βελτιστοποίηση και τι στερείται τώρα. Επίσης θα μπορεί να διαβάσει και να κατανοήσει τα δεδομένα και να τα ερμηνεύσει.
- Πολλοί επαγγελματίες έχουν διασυνδέσεις με άλλους ανθρώπους σε συγκεκριμένα πεδία, ειδήμονες σε θέσεις όπως δημοσιογράφους, διανομείς άρθρων, καταλόγων κτλ, που μπορούν να βοηθήσουν στην επιτυχία της επένδυσης για βελτιστοποίηση. Υπάρχουν και αναλυτές SEO που γνωρίζουν άτομα που εργάζονται σε μηχανές αναζήτησης αν και οι δεύτεροι δεν αποκαλύπτουν εύκολα τα μυστικά της μηχανής
- Οι επαγγελματίες έχουν την ικανότητα να προσδιορίσουν ποιες σελίδες είναι σχετικές με αυτή που βελτιστοποιούν και να γλυτώσουν χρόνο από το χτίσιμο των συνδέσμων. Επίσης είναι αρκετά ικανοί να αναγνωρίσουν ποιο κατάλογοι θα χρησιμοποιηθούν και που πιθανόν να δοθούν χρήματα για συνδέσμους (paid links)

- Επίσης καταφέρνουν να προσδιορίσουν και να χειριστούν θέματα που εμποδίζουν την σωστή ευρετηρίαση μιας ιστοσελίδας, της χαμηλής αξιολόγησης και των ποινών που πιθανόν επιβλήθηκαν από κάποια μηχανή αναζήτησης. Είναι μια ικανότητα που είναι δύσκολη στην απόκτηση της χωρίς πε-  
ρίρα χρόνων σε βελτιστοποιήσεις.
- Η εξοικονόμηση του χρόνου είναι ένα ατού στην επιλογή ενός επαγγελματία. Έχοντας όλο το θεωρητικό υπόβαθρο και τις καλές πρακτικές, εξοικονομείται χρόνος αναζήτησης γνώσης

### **Ατομική εργασία**

- Όταν ο χρήστης κάνει μόνος του την βελτιστοποίηση υπάρχει πλήρης έλεγχος της σελίδας. Ο χρήστης είναι γνώστης που βρίσκεται το κάθε τι, ποιος και τι δημιούργησε κάποιο σύνδεσμο ή τροποποιήθηκε κάτι.
- Με την εφαρμογή διαφόρων κανόνων ο χρήστης αποκτά εμπειρία και βλέπει μέσα από τις πράξεις του κατά πόσο καλά διδάχθηκε την βελτιστοποίηση. Οι αλλαγές και η συχνή παρακολούθηση της θέσης της σελίδα στα αποτελέσματα θα τον αξιολογήσει.
- Με αυτόν τον τρόπο, ο χρήστης αποκτά την πλήρη υπευθυνότητα. Η επιτυχία ή αποτυχία της βελτιστοποίησης θα κοστίζει στον ίδιο. Αυτό θα του δώσει κίνητρο να κάνει πιο σωστά την δουλεία του και δεν θα έχει κάποιο άλλοθι.
- Τέλος ο χρήστης εξοικονομεί χρηματικό ποσό που θα έδινε σε κάποιον επαγγελματία. Οι επενδύσεις τέτοιου είδους δεν είναι τόσο ασφαλείς γιατί στην αγορά υπάρχουν πολλοί που υποστηρίζουν ότι γνωρίζουν να βελτιστοποιούν, χωρίς να κάνουν εντρυφή στο θέμα.

## 4 Εργαλεία Βελτιστοποίησης

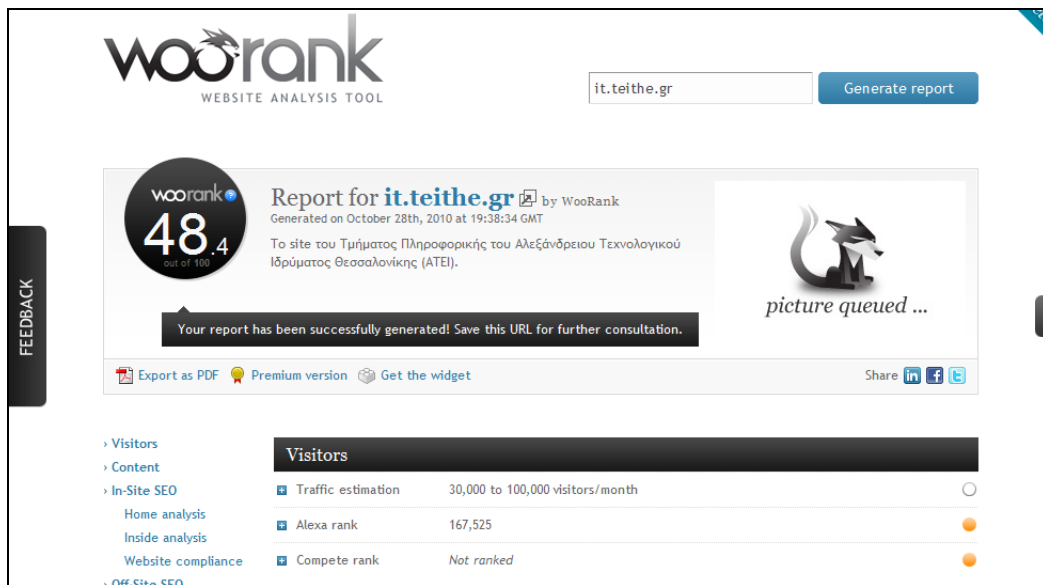
Τα εργαλεία βελτιστοποίησης είναι προγράμματα που εξετάζουν μια σελίδα σύμφωνα με τα κριτήρια βελτιστοποίησης που αναφέρθηκαν παραπάνω αλλά και βοηθητικά εργαλεία σχετικά με την βελτιστοποίηση. Τα εργαλεία αυτά δεν γράφουν μόνο τους κώδικα, ούτε εφαρμόζουν αλλαγές σε μια σελίδα αλλά ενημερώνουν τον χρήστη με αναφορές για διάφορα προβλήματα ή ενημερώσεις που θα πρέπει να γίνουν σε μια σελίδα για την καλύτερη βελτιστοποίηση.

Τα εργαλεία βελτιστοποίησης μπορεί να είναι γραμμένα σε Win32bit εφαρμογές αλλά και σε διαδικτυακές πλατφόρμες για ευκολία στην πρόσβαση και την αποτελεσματικότητά τους. Τα offline εργαλεία κυρίως απαιτούν σύνδεση με το Internet ώστε να εξετάζουν τις σελίδες αλλά και τους παράγοντες που την επηρεάζουν και ελάχιστα εξετάζουν σελίδες αποθηκευμένες στον δίσκο.

Το κόστος των εργαλείων ανέρχεται σε αρκετά ευρώ για την χρήση ενός εργαλείου ή επί πληρωμής πρόσβαση σε διάφορα εργαλεία ως υπηρεσία. Παλαιότερα, η βελτιστοποίηση ήταν κάπως άγνωστο σε πολλούς οπότε οι λίγοι που ήταν εξοικειωμένοι πίστευαν πολύ για τις υπηρεσίες τους. Πλέον, υπάρχει η τάση της δωρεάν χρήσης των εργαλείων από τους χρήστες ή η μηδαμινή χρέωση για την συντήρησή τους. Μηχανές αναζήτησης όπως η Google προσπαθούν να μπουν στο παιχνίδι παραχωρώντας οι ίδιες τους τα εργαλεία για βελτιστοποίηση αλλά και αποκαλύπτοντας τα μυστικά των αλγορίθμων τους ώστε να εντείνουν τον ανταγωνισμό.

### 4.1 Αναλυτές Βελτιστοποίησης

Είναι από τα πιο γνωστά εργαλεία που χρειάζεται όποιος θέλει να ασχοληθεί με την βελτιστοποίηση ιστοσελίδων. Τα εργαλεία αυτά σχεδιάστηκαν να αναλύουν τις σελίδες και να επιστρέφουν τα αποτελέσματα έτσι ώστε να ενημερωθεί ο χρήστης για την τωρινή κατάσταση.



Εικόνα 4.1. Ανάλυση ιστοσελίδας τμήματος Πληροφορικής στο Woorank.com

Για την χρήση αυτών των εργαλείων συνήθως ζητείται η διεύθυνση μιας σελίδας. Αφού ο χρήστης δώσει την επιθυμητή διεύθυνση, το εργαλείο διαβάζει τον HTML κώδικα και τον αναλύει σύμφωνα με τα κριτήρια που δόθηκαν στο 2<sup>ο</sup> κεφάλαιο. Μερικά εργαλεία ανάλυσης ζητούν ορισμένες λέξεις-κλειδιά ώστε να ελέγξουν την σχετικότητα των στοιχείων βελτιστοποίησης της σελίδας με αυτές που δόθηκαν ενώ άλλα προσπαθούν να εντοπίσουν το περιεχόμενο από την σελίδα.

Εξαρχής οι αναλυτές διαβάζουν και ελέγχουν τα μεταδεδομένα, ως προς το μέγεθος τους καθώς και τις λέξεις κλειδιά που περιέχουν. Μια άλλη λειτουργία τους είναι ο έλεγχος σχετικότητας του περιεχομένου της σελίδας με τα μεταδεδομένα. Έπειτα ακολουθεί μια ανάλυση συνδέσμων, εικόνων και λοιπών χαρακτηριστικών της σελίδας που αναφέρθηκαν στο 2<sup>ο</sup> κεφάλαιο.

Αφού ολοκληρωθεί ο έλεγχος της on-page βελτιστοποίησης ελέγχονται κάποια πιο γενικά στοιχεία για την ιστοσελίδα. Για παράδειγμα, ελέγχεται το domain και κάποιες πληροφορίες που το αφορούν, όπως η ηλικία του, στοιχεία όπως το Page Rank της σελίδας καθώς και ημερομηνίες όπου έγινε το τελευταίο crawling. Αντίστοιχο εργαλείο δημιουργήθηκε για την παρούσα πτυχιακή και αναλύεται παρακάτω. Μερικά online εργαλεία ανάλυσης είναι:

- <http://www.seoworkers.com/tools/analyzer.html>
- <http://websitegrader.com/>

- <http://www.reactionengine.com/>
- <http://spydermate.com/>
- <http://myseoanalyzer.com/>
- <http://www.woorank.com/>

## 4.2 Έρευνα και υπόδειξη λέξεων-κλειδιών

Ως γνωστόν, η αναζήτηση λέξεων-κλειδιών είναι το πρώτο βήμα που γίνεται όταν ξεκινά μια μελέτη βελτιστοποίησης για μια ιστοσελίδα. Αν αναλογιστεί κανείς ότι οι λέξεις-κλειδιά αποτελούν τη βάση της καμπάνιας, θα αντιληφθεί ότι η μελέτη και οι επιλογές θα πρέπει να γίνουν με μεγάλη προσοχή. Τα εργαλεία έρευνας και υπόδειξης λέξεων-κλειδιών σχεδιάστηκαν να ενημερώνουν τον χρήστη για συγκεκριμένες λέξεις-κλειδιά και να του προτείνουν σχετικές.

Λέξη-κλειδί	Αναγωνισμός	Συνολικές μηνιαίες αναζητήσεις	Τοπικές μηνιαίες αναζητήσεις	Τάσεις τοπικής αναζήτησης
joomla site		1.000.000	3.600	
dimiourgia istoselidas		18.100	18.100	
kataskeyh istoselidas		18.100	18.100	
δημιουργια ιστοσελιδας		18.100	18.100	
δημιουργια ιστοσελιδας		18.100	18.100	
kataskevi istoselidas		18.100	18.100	
κατασκευη ιστοσελιδας		18.100	18.100	
κατασκευη ιστοσελιδων		18.100	18.100	
κατασκευη ιστοσελιδων		18.100	14.800	
κατασκευη ιστοσελιδων		18.100	18.100	
κατασκευη ιστοσελιδων		18.100	14.800	
κατασκευη ιστοσελιδων		18.100	14.800	
κατασκευη ιστοσελιδων		18.100	14.800	
προωθηση		9.900	8.100	
ηλεκτρονικο εμποριο		5.400	5.400	
ηλεκτρονικα καταστηματα		4.400	4.400	
προωθηση ιστοσελιδων		4.400	4.400	

Εικόνα 4.2: Χρήση Google Ad words για την φράση " βελτιστοποίηση ιστοσελίδων"

Ο χρήστης δίνει μια λέξη-κλειδί που τον ενδιαφέρει και τα εργαλεία δημιουργούν μια αναφορά σχετικά με το πόσες φορές αναζητήθηκε η συγκεκριμένη λέξη-κλειδί στις δημοφιλείς μηχανές αναζήτησης τους προηγούμενους μήνες. Επίσης δίνονται νέες λέξεις-κλειδιά παρόμοιες με αυτές που δόθηκαν ως σημείο αναφοράς για νέες ιδέες. Αφού ολοκληρωθεί η έρευνα επιλέγονται ποιες λέξεις-κλειδιά περιγράφουν καλύτερα την σελίδα, ποιες είναι πιο ποιοτικές και ποιες αναζητούν οι χρήστες περισσότερο. Στην Εικόνα 4.2 απεικονίζεται ένα παράδειγμα αναφοράς σχετικών λέξεων-κλειδιών.

Η λειτουργία των εργαλείων βασίζεται στην αναζήτηση των όρων που δόθηκαν σε διάφορες μηχανές αναζήτησης και στην εξαγωγή συμπερασμάτων όπως ο αριθμός των αποτελεσμάτων ή οι σχετικοί όροι αναζήτησης. Με αυτό τον τρόπο συγκρίνουν τα αποτελέσματα της αναζήτησης με παρόμοιες αναζητήσεις και εμφανίζουν συγκεντρωτικά τα συμπεράσματα τους. Ορισμένα από αυτά είναι:

- <http://adwords.google.com/select/KeywordToolExternal>
- <http://www.webmaster-toolkit.com/keyword-research-tool.shtml>
- <http://www.digitalpoint.com/tools/suggestion/>
- <https://freekeywords.wordtracker.com/>

Τα εργαλεία αναφέρονται κυρίως στην αγγλική γλώσσα γιατί αναπτύχθηκαν από ξενόγλωσσους προγραμματιστές. Μοναδικό εργαλείο που υποστηρίζει την ελληνική γλώσσα είναι το εργαλείο υπόδειξης λέξεων-κλειδιών του Google Ad words.

### **4.3 Ανάλυση πυκνότητας λέξεων-κλειδιών**

Όπως έχει προαναφερθεί, η πυκνότητα των λέξεων-κλειδιών παίζει σημαντικό ρόλο στην βελτιστοποίηση μιας ιστοσελίδας, ρόλος που μπορεί εύκολα να ανατραπεί και να επηρεάσει αρνητικά την σελίδα. Γι αυτό πρέπει να δοθεί μεγάλη προσοχή στην συχνότητα επανάληψης των λέξεων-κλειδιών μέσα στο κείμενο.

Τα εργαλεία ανάλυσης της πυκνότητας επισκέπτονται μια ιστοσελίδα και εξαγουν μόνο το κείμενο της. Έπειτα, δημιουργούν ένα πίνακα με τις λέξεις που υπάρχουν στο κείμενο και καταμετρούν τις φορές που τις συναντούν στο κείμενο. Τέλος υπολογίζουν τον συνολικό αριθμό των λέξεων και βγάζουν τα ποσοστά που αντιστοιχούν σε κάθε λέξη. Ορισμένα εργαλεία δημιουργούν και ένα σύννεφο λέξεων κλειδιών, το λεγόμενο keyword cloud που παρουσιάζει οπτικά ποιες λέξεις-κλειδιά ξεχωρίζουν σε μια σελίδα. Υπενθυμίζεται ότι το κατάλληλο ποσοστό συχνότητας είναι περίπου στο 3-5%. Ορισμένα εργαλεία ανάλυσης της πυκνότητας των λέξεων κλειδιών βρίσκονται στις ακόλουθες διευθύνσεις:

- <http://tool.motoricerca.info/keyword-density.phtml>
- <http://tools.seobook.com/general/keyword-density/>
- <http://www.live-keyword-analysis.com/>



## 4.4 Υπολογιστής Page Rank

Έχουν δημιουργηθεί διάφορα εργαλεία, που ο χρήστης καλείται να δώσει την διεύθυνση μιας ιστοσελίδας και αυτά του επιστρέφουν το Page Rank. Τα εργαλεία δεν σχετίζονται με την Google και δεν αντικατοπτρίζουν την πραγματικότητα αλλά ύστερα από διάφορους υπολογισμούς δίνουν ένα εκτιμώμενο PR (Εικόνα 4.3).



Εικόνα 4.3: Πρόβλεψη PR από realprcheck.com

Τα εργαλεία υπολογισμού του PR δεν υπολογίζουν τον αλγόριθμο, διότι η Google δεν τον έχει κοινοποιήσει. Κάνουν διάφορους υπολογισμούς και ελέγχουν παράγοντες της σελίδας σχετικά με την Google ώστε να βρεθεί στο περίπου το PR. Πολλές φορές πέφτουν έξω ενώ άλλες φορές εμφανίζουν τον σωστό PR. Το μοναδικό εργαλείο που εμφανίζει το σωστό PR είναι το Google Toolbar, λόγω του ότι κατασκευάστηκε από την ομώνυμη μηχανή αναζήτησης. Μερικά εργαλεία πρόβλεψης Page Rank είναι:

- <http://www.prchecker.info/>
- <http://www.mypageranker.com/>
- <http://realprcheck.com/>
- <http://www.rustybrick.com/pagerank-prediction.php>
- [http://www.selfseo.com/google\\_pagerank\\_prediction\\_tool.php](http://www.selfseo.com/google_pagerank_prediction_tool.php)

## 4.5 Δημοτικότητα συνδέσμου

Link Popularity, ή αλλιώς δημοτικότητα συνδέσμου, ονομάζεται ο συνολικός αριθμός των ιστοσελίδων που συνδέονται σε μια δοσμένη ιστοσελίδα. Η δημοτικότητα μιας ιστοσελίδας μπορεί να αυξήσει δραματικά τους επισκέπτες της, γι αυτό και πολλοί χρήστες την ελέγχουν συχνά. Σύνδεσμοι τοποθετημένοι σε κάταλληλες ιστοσελίδες στέλνουν πιο στοχευμένο κοινό.

Περισσότερες από τις μεγαλύτερες μηχανές αναζήτησης υπολογίζουν τον παράγοντα αυτό ως παράμετρο στον αλγόριθμο σχετικότητας των ιστοσελίδων.

Ως αποτέλεσμα, όσο αυξάνονται οι ποιοτικοί σύνδεσμοι προς μια ιστοσελίδα, αυξάνεται και η κατάταξη της.

Τα εργαλεία που αναφέρονται παρακάτω ελέγχουν τις πιο σημαντικές μηχανές αναζήτησης, όπως Google, Yahoo!, Bing και αναφέρουν τον αριθμό των συνδέσεων ως προς την ιστοσελίδα σας. Κάθε μηχανή αναζήτησης προσφέρει την δυνατότητα στον χρήστη να ελέγχει αν κάποια ιστοσελίδα υπάρχει ως σύνδεσμος σε άλλους ιστοχώρους του ευρετηρίου της με την χρήση ορισμένων χαρακτηριστικών στην αναζήτηση, τα οποία αναφέρονται σε επόμενο κεφάλαιο. Τα εργαλεία δημοτικότητας συγκεντρώνουν τα αποτελέσματα και τα παρουσιάζουν στον χρήστη.

Domain	Total	Google	Yahoo	Live	AllTheWeb	AltaVista
teithe.gr	22410	292	9	22100	0	9
www.teithe.gr	11292	288	6084	1240	0	3680

Εικόνα 4.4: Έλεγχος δημοτικότητας από seocentro.com

Ορισμένα εργαλεία ελέγχου δημοτικότητας είναι:

- <http://www.seocentro.com/tools/search-engines/link-popularity.html>
- <http://www.marketleap.com/publinkpop/>
- <http://www.webmaster-toolkit.com/link-popularity-checker.shtml>
- <http://www.linkpopularity.com/>
- [http://www.iwebtool.com/link\\_popularity](http://www.iwebtool.com/link_popularity)
- <http://www.submitexpress.com/link-popularity-check.html>

## 4.6 Προσθήκη σε Καταλόγους

Υπάρχουν διάφορα εργαλεία που υπόσχονται να προσθέσουν οποιαδήποτε ιστοσελίδα τους δοθεί σε πάνω από 1,000 μηχανές αναζήτησης και καταλόγους. Τα εργαλεία αυτά σήμερα είναι αμφισβητούμενα και η χρήση τους έχει σχεδόν εξαλειφτεί.

Ως γνωστόν, οι σύνδεσμοι προς έναν ιστότοπο ήταν ένας από τους σημαντικότερους παράγοντες για την θέση του ιστότοπου σε μια μηχανή αναζήτησης. Την εποχή της άνθισης των φαρμών συνδέσμων ήταν δύσκολο κάποιος χρήστης να παρακολουθεί όλους τους καταλόγους που άνοιγαν και να καταχωρεί την ιστοσελίδα του.

Τα εργαλεία προώθησης σε καταλόγους κάνουν αυτήν ακριβώς την δουλειά. Συνήθως είναι σε online έκδοση αλλά συναντώνται και εκδόσεις για γνωστά λειτουργικά συστήματα. Τα εργαλεία αυτά έχουν μια αποθηκευμένη βάση με πολλούς καταλόγους και μηχανές αναζήτησης και ζητούν από τον χρήστη ορισμένα στοιχεία για την ιστοσελίδα του. Ανάλογα με τον κατάλογο ή την ιστοσελίδα, τα εργαλεία καταχωρούν τα δεδομένα από την σελίδα καταχώρησης ως πραγματικός χρήστης, δηλαδή θα συμπληρώσουν την φόρμα και θα την στείλουν.

Η τόσο γρήγορη και αλόγιστη καταχώρηση συνδέσμων σε ιστοτόπους μεγάλωσε το φαινόμενο του spamming. Πλέον, τα εργαλεία θεωρούνται άσκοπα λόγω της υποβάθμισης του παράγοντα inbound links (σύνδεσμοι προς μια ιστοσελίδα) και την ριζική αλλαγή του κανόνα. Επίσης, πολλοί κατάλογοι που διατηρούνται μέχρι σήμερα, έχουν προσθέσει κάποιο είδος προστασίας CAPTCHA (**C**ompletely **A**utomated **P**ublic **T**uring test to tell **C**omputers and **H**umans **A**part – Πλήρης Αυτοματοποιημένο Τεστ για διαχωρισμό των υπολογιστών από τους ανθρώπους) το οποίο θέτει τα εργαλεία προσθήκης σε μη αξιοποιήσιμα.

## 4.7 Εντολές αναζήτησης της Google

Η μηχανή αναζήτησης Google δίνει την δυνατότητα σε έναν χρήστη να ανακαλύψει διάφορα στοιχεία που αφορούν την σελίδα του σε σχέση με την Google. Μπορεί να βρει πόσες σελίδες της ιστοσελίδας τους έχουν ευρετηριαστεί, τι πληροφορίες έχει η Google και όχι μόνο. Ακολουθεί μια λίστα με τις σημαντικότερες εντολές.

- **site:**

Με την χρήση του [site:] στην αναζήτηση η Google περιορίζει τα αποτελέσματα της αναζήτησης στην διεύθυνση που δόθηκε. Με αυτό τον τρόπο ο χρήστης μπορεί να ελέγξει αν για κάποιος όρους εμφανίζεται η σελίδα του και πως εμφανίζεται στα αποτελέσματα [λουλούδια site:flower-shop.gr]

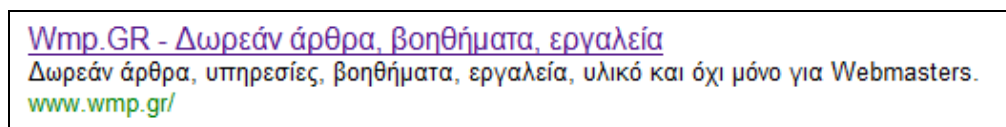
- **allintitle:**

Με την χρήση του [allintitle:] στην αναζήτηση η Google περιορίζει τα αποτελέσματα αναζήτησης σε σελίδες που περιέχουν στον τίτλο τους τις λέξεις που αναζητήθηκαν. [allintitle: πώληση ανθέων]

- intitle:**  
 Με την χρήση του [intitle:] στην αναζήτηση η Google περιορίζει τα αποτελέσματα σε σελίδες που έχουν τίτλο την πρώτη λέξη και τις υπόλοιπες στο σώμα κειμένου της. Για παράδειγμα το [intitle:πώληση λουλούδια] θα εμφανίζει αποτελέσματα σελίδων που περιέχουν την λέξη «πώληση» στον τίτλο της σελίδας και την λέξη «λουλούδια» σε οποιοδήποτε μέρος της.
- allinURL:**  
 Με την χρήση του [allinURL:] στην αναζήτηση η Google περιορίζει τα αποτελέσματα σε σελίδες που περιέχουν όλες τους όρους αναζήτησης στην διεύθυνση τους. Η εντολή δεν περιορίζει αν ο όρος βρίσκεται στο domain ή στην διαδρομή του αρχείου. Για παράδειγμα, χρησιμοποιώντας το [allinURL:flower shop] μπορεί να εμφανιστεί η σελίδα flower-shop.gr αλλά και σελίδες άσχετες όπως abc.gr/flower/shop, flower.gr/shop/abc
- inURL:**  
 Η χρήση της εντολής [inURL:] ενεργεί ακριβώς όπως η παραπάνω με την διαφορά ότι δέχεται μόνο την πρώτη λέξη να βρίσκεται στο domain και οι υπόλοιπες λέξεις να βρίσκονται στο κείμενο της σελίδας.
- cache:**  
 Με την χρήση της [cache:] στην αναζήτηση εμφανίζεται ποια έκδοση της σελίδας είναι αποθηκευμένη στην μηχανή αναζήτησης. Χρήσιμο σε αυτή τη αναζήτηση είναι η ημερομηνία που έγινε το τελευταίο caching.
- link:**  
 Η εντολή [link:] επιστρέφει όλες τις ιστοσελίδες που έχουν υπερσύνδεση προς την σελίδα που τέθηκε ως όρο στην αναζήτηση. Για παράδειγμα το [link:flower-shop.gr] θα εμφανίσει όλες τις ιστοσελίδες που δείχνουν στην flower-shop.gr
- related:**  
 Η εντολή [related:] επιστρέφει τις ιστοσελίδες που είναι παρόμοιες με την σελίδα που δόθηκε στην αναζήτηση. Με αυτό τον τρόπο εντοπίζονται ανταγωνιστές, παρόμοιες ιστοσελίδες αλλά και αν η μηχανή αναζήτησης κατανόησε το περιεχόμενο της ιστοσελίδας σωστά.

- **info:**

Με την χρήση της εντολής [info:] η Google επιστρέφει τις πληροφορίες που έχει για μια συγκριμένη ιστοσελίδα. Η [info:wmp.gr] θα επιστρέψει τις πληροφορίες που εμφανίζονται στην Εικόνα 4.5



Εικόνα 4.5: Αποτελέσματα εντολής info

Κάθε εντολή είναι μοναδική και παρέχει αρκετές πληροφορίες στον χρήστη ώστε να χρησιμοποιηθούν κατάλληλα για την βελτιστοποίηση της σελίδας του.

## 4.8 Seo-Browser.com

Ένα από τα πιο σημαντικά και δωρεάν εργαλεία είναι το SEO Browser. Το SEO browser είναι ένας εξομοιωτής για τον τρόπο που μια μηχανή αναζήτησης βλέπει τα δεδομένα μιας ιστοσελίδας και περιέχει και χρήσιμες πληροφορίες για το ότι συναντά. Ακολουθεί στιγμιότυπο της ιστοσελίδας του τμήματος Πληροφορικής στην Εικόνα 4.6.



Εικόνα 4.6: Προσομοίωση ιστοσελίδας www.it.teithe.gr

Η χρήση του SEO Browser είναι σχετικά εύκολη. Αρχικά, ο χρήστης επισκέπτεται την ηλεκτρονική διεύθυνση *www.seo-browser.com* και εισάγει την ηλεκτρονική διεύθυνση της δικής του ιστοσελίδας. Με την αποστολή της διεύθυνσης, βλέπει την ιστοσελίδα του όπως την βλέπει ο spider με πολύ λίγες πληροφορίες. Θα πρέπει να αλλάξει το mode από simple σε advanced για να λάβει όλες τις πληροφορίες που χρειάζεται. Με την επιλογή advanced αυτόματα ο χρήστης παίρνει μια γενική εικόνα για την ιστοσελίδα του. Ο SEO Browser δείχνει τον τίτλο της σελίδας και τα μεταδεδομένα, καθώς και στοιχεία όπως το βάρος σε KB της ιστοσελίδας, ο χρόνος φόρτωσης της. Σημαντικές πληροφορίες από το μπλοκ αυτό είναι η χρήση του robots.txt, ο χρόνος φόρτωσης και οι πληροφορίες για τις λέξεις του κειμένου.

Αμέσως μετά το πρώτο μπλοκ πληροφοριών, ο SEO Browser μας πληροφορεί για τις ανακατευθύνσεις (redirects) που δέχεται κάποιος επισκέπτης ή spider σε μια ιστοσελίδα. Αν η σελίδα δεν χρησιμοποιεί redirects, ο χρήστης θα δει HTTP/1.0 200 OK, ενώ αν χρησιμοποιεί κάποιο 301 ή 302 redirect θα ενημερωθεί κατάλληλα. Αυτός είναι ένας καλός τρόπος να δούμε αν η ιστοσελίδα έχει δεχτεί ή προβεί σε κακόβουλο SEO. Οι μηχανές αναζήτησης δέχονται θετικά μόνο το 301 redirect.

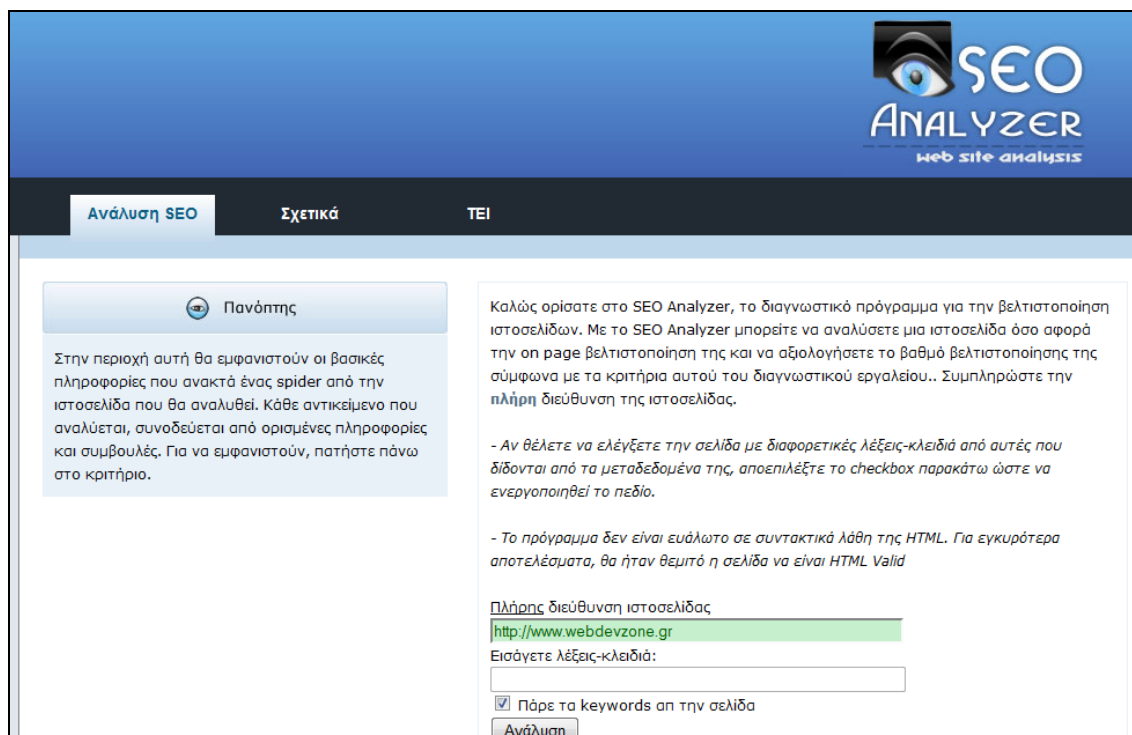
Αφού τελειώσουν όλα τα μπλοκ πληροφοριών ακολουθεί η ιστοσελίδα σε μορφή σειριακού κειμένου. Η πλοήγηση λειτουργεί κανονικά και ο χρήστης μπορεί να πλοηγηθεί στις υπόλοιπες σελίδες του ιστοτόπου μέσω του εξομοιωτή. Στο κείμενο παρατηρούνται ότι οι συνδέσεις με το χαρακτηριστικό «rel=nofollow» σημαίνονται από το εξομοιωτή με κόκκινη γραφή για τον διαχωρισμό τους από τις κανονικές συνδέσεις.

# 5 SEO Analyzer

Το SEO Analyzer είναι ένα εργαλείο ανάλυσης της δομής μιας ιστοσελίδας. Πρόκειται για ένα online εργαλείο, όπου ο χρήστης δίνει την διεύθυνση που επιθυμεί να εξετάσει και ο αναλυτής, αφού αναλύσει την σελίδα, του επιστρέφει μια σειρά από παρατηρήσεις και συμβουλές πάνω σε παράγοντες βελτιστοποίησης που εντοπίζονται στην σελίδα καθώς και μια γενική βαθμολογία για πόσο βελτιστοποιημένη είναι η σελίδα.

## 5.1 Γενική Χρήση

Η χρήση του εργαλείου είναι απλή. Ο χρήστης θα πρέπει να πλοηγηθεί στην διεύθυνση που έχει εγκαταστήσει το πρόγραμμα ανάλυσης στον δικό του server. Εναλλακτικά, μπορεί να πλοηγηθεί στην διεύθυνση που φιλοξενείται η εφαρμογή, <http://www.webdevzone.gr/ergalia/seo-analyzer.html>. Η αρχική εικόνα που συναντά ο χρήστης, εμφανίζεται στην εικόνα Εικόνα 5.1.



**SEO ANALYZER**  
web site analysis

Ανάλυση SEO    Σχετικά    ΤΕΙ

**Πανόπτης**

Στην περιοχή αυτή θα εμφανιστούν οι βασικές πληροφορίες που ανακτά ένας spider από την ιστοσελίδα που θα αναλυθεί. Κάθε αντικείμενο που αναλύεται, συνοδεύεται από ορισμένες πληροφορίες και συμβουλές. Για να εμφανιστούν, πατήστε πάνω στο κριτήριο.

Καλώς ορίσατε στο SEO Analyzer, το διαγνωστικό πρόγραμμα για την βελτιστοποίηση ιστοσελίδων. Με το SEO Analyzer μπορείτε να αναλύσετε μια ιστοσελίδα όσο αφορά την on page βελτιστοποίηση της και να αξιολογήσετε το βαθμό βελτιστοποίησης της σύμφωνα με τα κριτήρια αυτού του διαγνωστικού εργαλείου.. Συμπληρώστε την πλήρη διεύθυνση της ιστοσελίδας.

- Αν θέλετε να ελέγξετε την σελίδα με διαφορετικές λέξεις-κλειδιά από αυτές που δίδονται από τα μεταδεδομένα της, αποπιλέξτε το checkbox παρακάτω ώστε να ενεργοποιηθεί το πεδίο.

- Το πρόγραμμα δεν είναι ευάλωτο σε συντακτικά λάθη της HTML. Για εγκυρότερα αποτελέσματα, θα ήταν θεμιτό η σελίδα να είναι HTML Valid

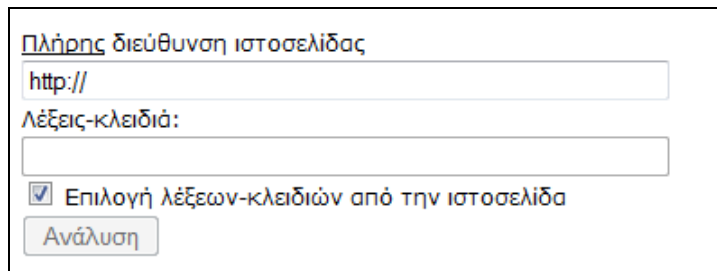
Πλήρης διεύθυνση ιστοσελίδας

Εισάγετε λέξεις-κλειδιά:

Πάρε τα keywords απ την σελίδα

Εικόνα 5.1: Αρχική σελίδα εφαρμογής

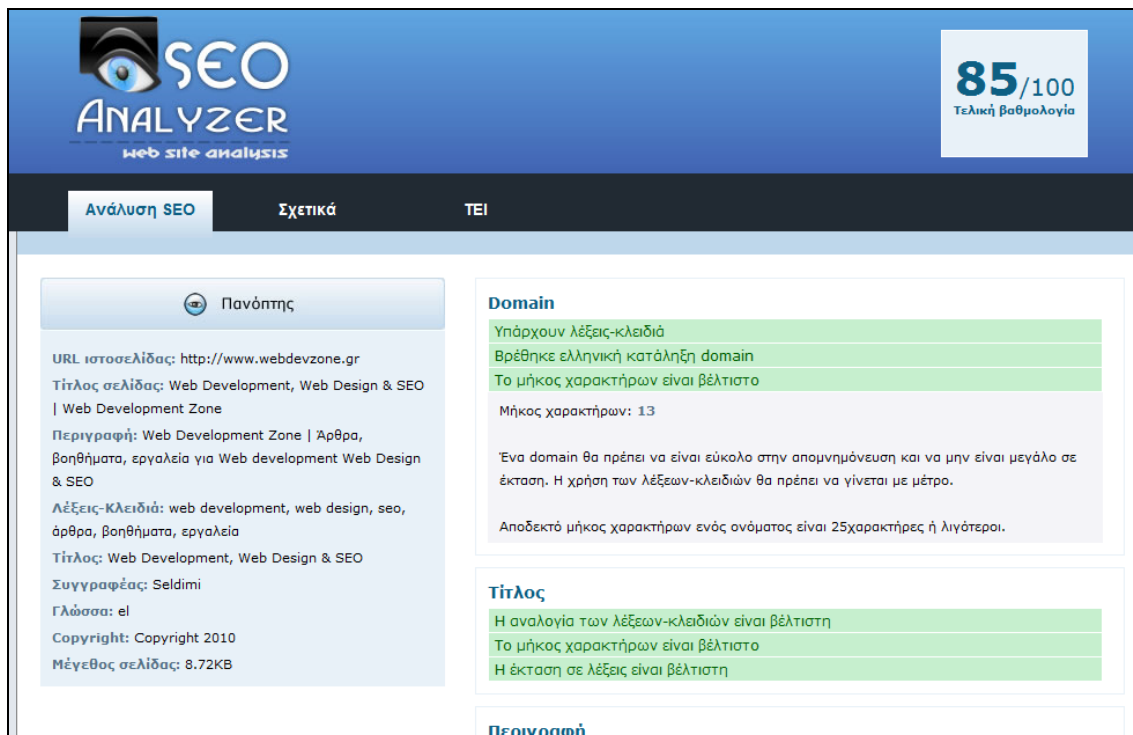
Για την σωστή λειτουργία του προγράμματος, θα πρέπει να δοθούν ορισμένες λέξεις-κλειδιά. Ο χρήστης έχει την δυνατότητα να ανακτήσει τις λέξεις-κλειδιά από την σελίδα που εξετάζει, από τα μεταδεδομένα, ή αλλιώς μπορεί να εξετάσει λέξεις ή φράσεις που επιθυμεί αφαιρώντας την επιλογή «Επιλογή λέξεων-κλειδιών από την ιστοσελίδα», όπως φαίνεται στην Εικόνα 5.2, και να πειραματιστεί.



Πλήρης διεύθυνση ιστοσελίδας  
http://  
Λέξεις-κλειδιά:  
 Επιλογή λέξεων-κλειδιών από την ιστοσελίδα  
Ανάλυση

Εικόνα 5.2: Αρχική οθόνη SEO Analyzer

Όταν τα στοιχεία της φόρμας συμπληρωθούν ο χρήστης πατάει το κουμπί «Ανάλυση». Ο αναλυτής επισκέπτεται την ιστοσελίδα και εξάγει κάποια αποτελέσματα. Μόλις ολοκληρωθεί η διαδικασία, παρουσιάζεται στον χρήστη μια αναφορά από την ανάλυση και η βαθμολογία του, όπως φαίνεται στην Εικόνα 5.3.



**SEO ANALYZER**  
web site analysis

85/100  
Τελική βαθμολογία

Ανάλυση SEO    Σχετικά    ΤΕΙ

🌐 Πανόπτης

URL ιστοσελίδας: http://www.webdevzone.gr  
Τίτλος σελίδας: Web Development, Web Design & SEO | Web Development Zone  
Περιγραφή: Web Development Zone | Άρθρα, βοηθήματα, εργαλεία για Web development Web Design & SEO  
Λέξεις-Κλειδιά: web development, web design, seo, άρθρα, βοηθήματα, εργαλεία  
Τίτλος: Web Development, Web Design & SEO  
Συγγραφέας: Seldimi  
Γλώσσα: el  
Copyright: Copyright 2010  
Μέγεθος σελίδας: 8.72KB

**Domain**  
Υπάρχουν λέξεις-κλειδιά  
Βρέθηκε ελληνική κατάληξη domain  
Το μήκος χαρακτήρων είναι βέλτιστο  
Μήκος χαρακτήρων: 13  
Ένα domain θα πρέπει να είναι εύκολο στην απομνημόνευση και να μην είναι μεγάλο σε έκταση. Η χρήση των λέξεων-κλειδιών θα πρέπει να γίνεται με μέτρο.  
Αποδεκτό μήκος χαρακτήρων ενός ονόματος είναι 25χαρακτήρες ή λιγότεροι.

**Τίτλος**  
Η αναλογία των λέξεων-κλειδιών είναι βέλτιστη  
Το μήκος χαρακτήρων είναι βέλτιστο  
Η έκταση σε λέξεις είναι βέλτιστη

**Περιγραφή**

Εικόνα 5.3: Αναφορά ανάλυσης



Πάνω δεξιά εμφανίζεται η βαθμολογία της σελίδας και ακολουθούν αριστερά οι πληροφορίες που αντλήθηκαν από τα μεταδεδομένα και στο κύριο μέρος της σελίδας τα χαρακτηριστικά που αναλύθηκαν, χωρισμένα με κατηγορίες, και τρόπους βελτίωσης τους. Κάθε χαρακτηριστικό παίρνει ένα χρώμα, ανάλογα με το αν επηρεάζει θετικά την ιστοσελίδα ή αρνητικά στον τομέα της βελτιστοποίησης. Τα χαρακτηριστικά με κόκκινο χρώμα επιδρούν αρνητικά για την σελίδα, ενώ τα πράσινα κριτήρια είναι θετικά. Όσα χαρακτηριστικά έχουν κίτρινο χρώμα θεωρούνται ως θετικοί παράγοντες αλλά όχι σε μέγιστο βαθμό. Με λίγα λόγια, το κίτρινο χρώμα εκφράζει ένα ικανοποιητικό επίπεδο που έχει περιθώρια βελτίωσης ενώ το κόκκινο σημεία που χρήζουν άμεσης τροποποίησης.

Κάθε χαρακτηριστικό συνοδεύεται από επεξηγηματικό κείμενο με μια μικρή αναφορά στον τρόπο που επηρεάζει καθώς και ορισμένες συμβουλές βελτίωσης. Για την ανάγνωση του συνοδευτικού κειμένου, ο χρήστης πρέπει να πατήσει πάνω σε ένα χαρακτηριστικό. Το συνοδευτικό κείμενο αναδιπλώνει, ανάλογα με τις επιθυμίες του χρήστη.

The screenshot displays the 'SEO Analyzer' interface. At the top left is the logo 'SEO ANALYZER web site analysis'. At the top right, a box shows the score '85/100' with the text 'Τελική βαθμολογία'. Below the header is a navigation bar with 'Ανάλυση SEO', 'Σχετικά', and 'ΤΕΙ'. The main content area is divided into two columns. The left column, titled 'Πανόπτης', lists meta-information: URL (http://www.webdevzone.gr), Title (Web Development, Web Design & SEO | Web Development Zone), Description (Web Development Zone | Άρθρα, βοηθήματα, εργαλεία για Web development Web Design & SEO), Keywords (web development, web design, seo, άρθρα, βοηθήματα, εργαλεία), Title (Web Development, Web Design & SEO), Author (Seldimi), Language (el), Copyright (Copyright 2010), and Page Size (8.72KB). The right column contains analysis results for 'Domain', 'Title', and 'Description'. The 'Domain' section shows 'Υπάρχουν λέξεις-κλειδιά', 'Βρέθηκε ελληνική κατάληξη domain', and 'Το μήκος χαρακτήρων είναι βέλτιστο'. The 'Title' section shows 'Η αναλογία των λέξεων-κλειδιών είναι βέλτιστη', 'Το μήκος χαρακτήρων είναι βέλτιστο', and 'Η έκταση σε λέξεις είναι βέλτιστη'. The 'Description' section is partially visible.

Εικόνα 5.4: Αποτελέσματα ανάλυσης

## 5.2 Τεχνική Περιγραφή

Ο SEO Analyzer είναι προγραμματισμένος σε γλώσσα PHP και τρέχει σε web server apache. Για την λειτουργία του χρησιμοποιήθηκε ο *PHP Simple HTML DOM Parser* για την ανάκτηση του DOM και των αντικειμένων της σελίδας και ορισμένες έτοιμες κλάσεις ενώ για τις ανάγκες της διεπαφής χρησιμοποιήθηκε το πρόγραμμα Adobe Fireworks για τα γραφικά και η γλώσσα HTML με συνοδεία CSS. Τέλος, χρησιμοποιήθηκε η βιβλιοθήκη jQuery για την αναδίπλωση του συνοδευτικού κειμένου.

Για τις ανάγκες της επίδειξης χρησιμοποιήθηκε ο apache. Ο apache είναι ένας web server ανοιχτού λογισμικού. Είναι ο πιο διαδεδομένος στον χώρο των ιστοσελίδων και συνοδεύει κυρίως συστήματα βασισμένα σε Unix/Linux. Υποστηρίζει πολλές γλώσσες προγραμματισμού, με κυριότερη την PHP. Ο SEO Analyzer τρέχει και σε άλλους servers όπως ο Microsoft IIS, lighthttpd. Δεν απαιτεί κάποια ξεχωριστή βιβλιοθήκη της PHP αλλά χρησιμοποιεί βασικές κλήσεις μεθόδων.

Για την εξέταση των χαρακτηριστικών που έχει η ιστοσελίδα γίνεται χρήση του *PHP Simple HTML DOM Parser* (<http://simplehtmldom.sourceforge.net/>), ενός open source parser για την ανάκτηση της html και του DOM μιας ιστοσελίδας. Ο parser επισκέπτεται την σελίδα που δίνει ο χρήστης και δημιουργεί πίνακες (arrays) με στοιχεία που αφορούν την δομή του DOM. Ο SEO Analyzer αναλύει τα δεδομένα του DOM εφαρμόζοντας διάφορα κριτήρια μέτρησης και αντιστοίχισης με την χρήση κανονικών εκφράσεων ώστε να εξαχθούν τα συμπεράσματα.

Η χρήση του jQuery γίνεται στην οπτική διεπαφή του χρήστη με τον SEO analyzer. Το jQuery είναι μια βιβλιοθήκη ανοιχτού κώδικα, βασισμένη στην JavaScript. Με την χρήση του jQuery ο χρήστης μπορεί να γράψει πιο κατανοητό και καθαρό κώδικα αλλά και να αυτοματοποιήσει διαδικασίες. Συγκεκριμένα, ο SEO Analyzer χρησιμοποιεί της μεθόδους `.show()` και `.hide()` για την αναδίπλωση του συνοδευτικού κειμένου.

Τέλος γίνεται η χρήση μιας κλάσης ανοιχτού κώδικα της wordCloud. Ο προγραμματιστής της κλάσης είναι ο Derek Harvey και η κλάση βρίσκεται στην ιστοσελίδα [www.lotsofcode.com](http://www.lotsofcode.com) Η κλάση δημιουργεί ένα αντικείμενο cloud και

δέχεται ως όρισμα ένα κείμενο. Η κλάση δέχεται το κείμενο και μετράει πόσες φορές εντοπίζεται η κάθε λέξη στο κείμενο. Στο τέλος εξάγει τις λέξεις του κειμένου με τυχαία σειρά και με μέγεθος γραφής σχετικό της αναλογίας της κάθε λέξης. Με αυτόν τον τρόπο επιτυγχάνεται η μέτρηση της πυκνότητας των λέξεων στο κείμενο της σελίδας, βασικό στοιχείο βελτιστοποίησης.

### 5.3 Τεχνικές Δυσκολίες

Δυο σημαντικά θέματα προέκυψαν κατά την υλοποίηση του προγράμματος ανάλυσης, θέματα κυρίως προγραμματιστικά αλλά και την αποτυχία των μηχανών να αντιληφθούν τον γραπτό λόγο και τις ιδιαιτερότητες της κάθε γλώσσας.

Η πρώτη δυσκολία εμφανίζεται όταν στο πρόγραμμα είναι απαραίτητο να χρησιμοποιηθούν οι έτοιμες συναρτήσεις της PHP. Δυστυχώς η PHP δεν υποστηρίζει πλήρως utf-8, utf-16 αλλά ISO κωδικοποιήσεις στις μεθόδους διαχείρισης strings και arrays. Αυτό έχει ως αποτέλεσμα μεθόδους όπως η strtolower(), μέθοδος αλλαγής γραφής σε πεζά γράμματα, ή η strlen(), μέθοδος μέτρησης χαρακτήρων, να παραμορφώνουν το αρχικό string ή να επιστρέφουν λάθος αποτελέσματα.

Λύση στο πρόβλημα είναι η χρήση της utf8\_decode(), μιας μεθόδου που προσπαθεί να αποκωδικοποιήσει σε utf-8 μορφή αλλά δεν λειτουργεί πάντα επιθυμητά. Πολλές φορές η ελληνικές λέξεις μετατρέπονται σε ερωτηματικά. Καλύτερη λύση αποτελεί η δημιουργία παρόμοιων μεθόδων και υλοποίηση τους με κανονικές εκφράσεις ή συνδυασμό άλλων μεθόδων. Με αυτόν τον τρόπο αντιμετωπίζονται επιτυχώς όλες οι ιδιαιτερότητες μιας μεθόδου.

Η δεύτερη δυσκολία είναι η χρήση της ελληνικής γλώσσας και των κανονικών εκφράσεων. Οι κανονικές εκφράσεις χρησιμοποιούν κυρίως ASCII χαρακτήρες ενώ νεότερες μηχανές υιοθετούν κωδικοποιήσεις utf-8. Κάθε ελληνικός χαρακτήρας αντιστοιχίζεται με μια ascii τιμή ώστε η μηχανή να μπορεί να τον αντιληφθεί και να τον χρησιμοποιήσει. Το πρόβλημα συναντάται στην χρήση ορισμένων κλάσεων των κανονικών εκφράσεων. Για παράδειγμα η έκφραση [w]+ αντιστοιχεί σε μια λέξη. Η κλάση αυτή όμως υποστηρίζει μόνο τους αγγλικούς χαρακτήρες. Ως αποτέλεσμα, οι ελληνικές λέξεις να μην μπορούν να αντιστοιχιστούν και η κανονική έκφραση να δυσλειτουργεί.

Λύση στο πρόβλημα είναι η χρήση της έκφρασης ζεύγους χαρακτήρων [Α-ώ]+ ώστε να παρθεί η ascii τιμή του τονισμένου κεφαλαίου γράμματος Α που αποτελεί αρχή του ελληνικού αλφαβήτου για έναν Η/Υ και το τονισμένο γράμμα ω που είναι το τελευταίο. Οι ενδιάμεσες τιμές περιέχουν όλο το ελληνικό αλφάβητο, τονισμένο και μη καθώς και τα φωνήεντα με διαλυτικά.

## 6 Συμπεράσματα

Οι ιστοσελίδες αποτελούν το σημαντικότερο μέσο ενημέρωσης και διακίνησης πληροφοριών της εποχής μας. Ανεξάρτητα για το αν πρόκειται για ένα εμπορικό site ή ένα προσωπικό blog, η επιθυμία να προσελκύσει η σελίδα επισκέπτες είναι δεδομένη στους ιδιοκτήτες των σελίδων. Η βελτιστοποίηση ιστοσελίδων ως προς τις μηχανές αναζήτησης είναι μια σύνθετη διαδικασία βελτίωσης μιας ιστοσελίδας που αποσκοπεί στο να πριμοδοτούνται οι συγκεκριμένες ιστοσελίδες από τις μηχανές αναζήτησης και να προβάλλονται υψηλότερα στα αποτελέσματα αναζήτησης. Ο ανταγωνισμός είναι μεγάλος, γιατί οι σελίδες που κατατάσσονται στην κορυφή των αποτελεσμάτων δέχονται τους περισσότερους επισκέπτες.

Ένα άλλο ενδιαφέρον στοιχείο που παρατηρεί όποιος ασχολείται με την βελτιστοποίηση είναι το άγνωστο και η πρόκληση στην εξερεύνηση του. Ως γνωστόν, οι αλγόριθμοι αναζήτησης είναι κρυφοί και αλλάζουν συχνά. Ο καθένας δημιουργεί τις δικές του εικασίες και πειραματίζεται. Η αποκρυπτογράφηση των αλγορίθμων είναι όπως η ανάγκη του ανθρώπου να γνωρίσει την φύση. Με πειράματα και εικασίες καταλήγει στο πως, από τι και γιατί γίνεται κάτι. Στην βελτιστοποίηση ιστοσελίδων ο χρήστης δημιουργεί τα δικά του θεωρήματα, τα εφαρμόζει σε διάφορες ιστοσελίδες και παρακολουθεί τα αποτελέσματα. Αν επιβεβαιωθεί, μπορεί να το αναπαράγει σε άλλες ιστοσελίδες ή να το τροποποιήσει ώστε να πετύχει κάτι καλύτερο. Όταν εικασίες επιβεβαιωθούν από πολλούς, μπορούν να θεωρηθούν ως τεχνική βελτιστοποίησης. Ο χρόνος επιβεβαίωσης όμως είναι το μειονέκτημα, μιας που τα αποτελέσματα μπορεί να εμφανιστούν μετά από μήνες εργασίας και παρακολούθησης.

Το πρόβλημα στην πτυχιακή ήταν η εύρεση των στοιχείων που επηρεάζουν θετικά μια ιστοσελίδα. Όπως προαναφέρθηκε, τα περισσότερα αποτελούν εικασίες χρηστών που επιβεβαιώθηκαν. Για να γίνει έγκυρη η πτυχιακή, καταγράφηκαν τα στοιχεία που επιβεβαιώθηκαν από τους περισσότερους αναλυτές αλλά και από ανθρώπους που δουλεύουν σε μηχανές αναζήτησης και δίνουν

συμβουλές για το πως θα πρέπει να δομηθεί το περιεχόμενο μιας σελίδας, όπως πρόσφατα έκανε η Google στο «Webmaster's Central Blog».<sup>3</sup>

Το πρόγραμμα που αναπτύχθηκε στα πλαίσια της πτυχιακής αυτοματοποιεί το μεγαλύτερο μέρος της μελέτης των στοιχείων της ιστοσελίδας. Το πρόγραμμα είναι ικανό να βοηθήσει τον χρήστη να αντιληφθεί τι είναι σημαντικό για την σελίδα του, πως είναι η τρέχουσα κατάσταση και τι βελτιώσεις θα μπορέσει να εφαρμόσει. Το πρόγραμμα, όπως και η βελτιστοποίηση ως ενέργεια, δεν αυτοματοποιούνται καθώς πολλά ζητήματα απαιτούν ανθρώπινη παρέμβαση και δεν μπορεί να αυτοματοποιηθούν. Παρόλα αυτά, η εφαρμογή μπορεί χρησιμοποιηθεί σε ένα site για διάγνωση μια σειράς "ζητημάτων" και παροχή συμβουλών για την αντιμετώπισή τους.

Το πρόγραμμα επικεντρώνεται στην διάγνωση των στοιχείων που επηρεάζουν την βελτιστοποίησης. Ο χρήστης μπορεί να διαγνώσει τα στοιχεία που επηρεάζουν την αξιολόγηση της σελίδας και να τα τροποποιήσει. Η παρακολούθηση της προόδου θα πρέπει να γίνει με εξωτερικά εργαλεία όπως το Google analytics<sup>4</sup> ή να αναπτυχθεί, ως μελλοντική εργασία, ένα εργαλείο συγκέντρωσης δεδομένων από τις μηχανές αναζήτησης. Τα εργαλεία συγκέντρωσης δεδομένων χρησιμοποιούν parsers και λειτουργούν διαβάζοντας την html από τα αποτελέσματα αναζήτησης από εντολές όπως η *link:istoselida.gr*. Τα εργαλεία χρειάζονται συχνό έλεγχο και ανανέωση γιατί η μηχανή αναζήτησης μπορεί να αλλάξει τον html κώδικα της ανά πάσα στιγμή.

Το μειονέκτημα της εφαρμογής, αλλά και της πτυχιακής είναι οι συνεχείς αλλαγές, όπως αναφέρθηκε και στον πρόλογο. Όσο αυξάνονται οι απαιτήσεις για την εύρεση της πιο έγκυρης και σχετικής πληροφορίας, τόσο θα διαμορφώνονται τα στοιχεία που επηρεάζουν την κατάταξη μιας σελίδας. Πολύ πιθανόν, σε ορισμένο διάστημα, οι πληροφορίες της πτυχιακής να είναι ξεπερασμένες, ή μην ισχύουν για ορισμένα στοιχεία.

Η παρούσα πτυχιακή θα μπορούσε να επεκταθεί με την ενσωμάτωση στην εφαρμογή ικανότητας ελέγχου συμμόρφωσης με περισσότερους "κανόνες" βελτιστοποίησης. Επιπλέον, ενδιαφέρον θα είχε η ικανότητα από μέρους της

---

<sup>3</sup> Official Google Webmaster Central Blog – <http://googlewebmastercentral.blogspot.com/>

<sup>4</sup> Google Analytics - <http://www.google.com/analytics/>

εφαρμογής να διαχειρίζεται ως προς τα ζητήματα SEO έναν πλήρη ιστοτόπο και όχι αποσπασματικά (και ανεξάρτητα) τις σελίδες του.

Εν τέλει, η βελτιστοποίηση των ιστοσελίδων θα πρέπει να λαμβάνεται υπόψη σε όλες τις φάσεις της ζωής τους, τόσο στη σχεδίαση και ανάπτυξη όσο και στη λειτουργία και τη συντήρησή τους.





# Αναφορές

1. Bombay Cow: <http://www.bombaycrow.com> (Ανάκτηση: 17 Σεπτεμβρίου 2010)
2. CLCS.net, *Page Hijack: The 302 Exploit, Redirects and Google*, <http://clsc.net/2005/internet/google-302-page-hijack/> (Ανάκτηση: 17 Σεπτεμβρίου 2010)
3. ClickMedia, *Τι είναι το Keyword Density*, <http://www.click-media.gr/knowledge-base/what-is-keyword-density.html> (Ανάκτηση: 17 Σεπτεμβρίου 2010)
4. Direct Traffic Media, *Off Page Optimization VS On Page Optimization*, [http://www.directtrafficmedia.co.uk/on\\_page\\_optimisation.htm](http://www.directtrafficmedia.co.uk/on_page_optimisation.htm) (Ανάκτηση: 17 Σεπτεμβρίου 2010)
5. Google Webmaster Central Blog, <http://googlewebmastercentral.blogspot.com/> (Ανάκτηση: 17 Σεπτεμβρίου 2010)
6. Hostsun, *Ρόλος του Google Page Rank (βαθμολόγηση ιστοσελίδων) για καλή θέση στα αποτελέσματα των μηχανών αναζήτησης*, <http://hostsun.gr/website-seo-google-page-rank.html> (Ανάκτηση: 17 Σεπτεμβρίου 2010)
7. Ogranic SEO, *Choosing Your Target Keywords*, <http://www.organicseo.org/book/keywords>, (Ανάκτηση: 17 Σεπτεμβρίου 2010)
8. Search Engine Journal, <http://www.searchenginejournal.com>, (Ανάκτηση: 17 Σεπτεμβρίου 2010)
9. SEO Book, <http://www.seobook.com>, (Ανάκτηση: 17 Σεπτεμβρίου 2010)
10. SEO Chat, *Domains and SEO*, <http://www.seochat.com/c/a/Search-Engine-Optimization-Help/Domains-and-SEO/> (Ανάκτηση: 17 Σεπτεμβρίου 2010)

11. Shimon Sandler, *14 Methods to Hide Text* ,  
<http://www.shimonsandler.com/14-methods-to-hide-text/> (Ανάκτηση: 17 Σεπτεμβρίου 2010)
12. Strategy Internet Marketing, *PPC or SEO - Which Should You Choose*,  
<http://www.strategyinternetmarketing.co.uk/ppc-or-seo.html> (Ανάκτηση: 17 Σεπτεμβρίου 2010)
13. Web Design Library, *SEO: Real Black Hat Secrets Unleashed and How to Avoid it*, <http://www.webdesign.org/site-maintenance/web-promotion/seo-real-black-hat-secrets-unleashed-and-how-to-avoid-it.12895.html>  
(Ανάκτηση: 17 Σεπτεμβρίου 2010)
14. WebDesignForums.NET, <http://www.webdesignforums.net> (Ανάκτηση: 17 Σεπτεμβρίου 2010)
15. Wikipedia: *Search Engine*,  
[http://en.wikipedia.org/wiki/Web\\_search\\_engine](http://en.wikipedia.org/wiki/Web_search_engine) (Ανάκτηση: 17 Σεπτεμβρίου 2010)
16. Webdev Zone, <http://www.webdevzone.gr> (Ανάκτηση: 17 Σεπτεμβρίου 2010)
17. WebSeo Analytics, <http://www.webseoanalytics.com> (Ανάκτηση: 17 Σεπτεμβρίου 2010)

# Βιβλιογραφία

1. Chaffey D. (2007). *Search Engine Optimization Best Practice Guide 2007*. E-consultancy.com
2. Cristian Darie C. and Sirovich J. (2007). *Professional Search Engine Optimization with PHP*. Wiley Publishing, Inc.
3. Jennifer Grappone J. and Couzin G. (2008), *Search Engine Optimization An Hour a Day, second edition*. Wiley Publishing, Inc.
4. Ledford J.L., (2007). *Search Engine Optimization Bible*. Wiley Publishing, Inc
5. Wall M. A., (2005). *Search Engine Optimization Book*



# Παράρτημα

## Άλλοι Αλγόριθμοι Κατάταξης Ιστοσελίδων

### TrustRank

Το TrustRank είναι μια τεχνική ανάλυσης συνδέσμων, που δημιουργήθηκε από τον Stanford University και τους ερευνητές της Yahoo! Για ημιαυτόματο διαχωρισμό των χρήσιμων ιστοσελίδων από το spam.

Πολλές spam ιστοσελίδες δημιουργούνται με σκοπό την παραπλάνηση των μηχανών αναζήτησης. Οι spam σελίδες δημιουργούνται κυρίως για εμπορικούς σκοπούς και χρησιμοποιούν διάφορες τεχνικές για να επιτύχουν υψηλότερες κατατάξεις απ' ότι πραγματικά αξίζουν στα αποτελέσματα αναζητήσεων. Εμπειρογνώμονες μπορούν εύκολα να εντοπίσουν spam ιστοσελίδες, αλλά το κόστος αξιολόγησης τεράστιο αριθμού ιστοσελίδων είναι απαγορευτικό

Η μέθοδος TrustRank απαιτεί την επιλογή ενός μικρού συνόλου σελίδων που αξιολογούνται από έναν εμπειρογνώμονα. Το σύνολο των σελίδων χαρακτηρίζονται ως σελίδες-σπόροι και αποτελούν σελίδες χρήσιμες και αξιόπιστες. Με αυτό τον τρόπο «μαθαίνουν» σε μια μηχανή αναζήτησης τι χαρακτηριστικά έχει μια αξιόπιστη σελίδα.. Τη στιγμή που οι αξιόπιστες σελίδες θα αναγνωριστούν, ένας crawler επεκτείνεται έξω από το σύνολο των σπόρων αναζητώντας παρόμοιες σελίδες ή ακλουθώντας υπερσυνδέσεις από αυτούς. Η αξιοπιστία του TrustRank εξασθενεί καθώς οι σελίδες απομακρύνονται από τις σελίδες-σπόρους.

### HITS

Ο *Hyperlink-Induced Topic Search (HITS)*, γνωστός και ως *Hubs and authorities*, είναι ένας αλγόριθμος ανάλυσης συνδέσμων που αξιολογεί ιστοσελίδες. Ο αλγόριθμος αναπτύχθηκε από Jon Kleinberg. Ο αλγόριθμος καθορίζει δύο τιμές για κάθε σελίδα, την αξία του περιεχομένου της σελίδας και την αξία των υπερσυνδέσεων με άλλες σελίδες.

Στον αλγόριθμο HITS, το πρώτο βήμα είναι η ανάκτηση ενός συνόλου αποτελεσμάτων στο ερώτημα αναζήτησης. Ο υπολογισμός γίνεται μόνο στο συγκεκριμένο σύνολο αποτελεσμάτων και όχι σε όλες τις ιστοσελίδες.

Οι αξίες του περιεχομένου της σελίδας και των υπερσυνδέσεων ορίζονται με αμοιβαία. Η αξία του περιεχόμενου ορίζεται ως το σύνολο των κλιμακωτών συνδέσεων που οδηγούν σε αυτή τη σελίδα. Η αξία υπερσυνδέσεων ορίζεται ως το σύνολο των κλιμακωτών αξιών περιεχομένου που η σελίδα οδηγεί. Κάποιες εφαρμογές εξετάζουν την σχετικότητα των συνδεδεμένων σελίδων.

Ο υπολογισμός γίνεται με τις ακόλουθες αναδρομές. Κάθε σελίδα έχει αξία περιεχομένου και συνδέσμων ίση με 1. Η ιστοσελίδα παίρνει αξία περιεχομένου με το να συνδέονται σε αυτήν ιστοσελίδες που αναγνωρίζονται ως σελίδες πληροφοριών, δηλαδή σελίδες που έχουν υψηλή αξία συνδέσμων. Το αντίστροφο συμβαίνει με την αξία συνδέσμων. Η σελίδα παίρνει αξία υπερσυνδέσεων όταν οι σύνδεσμοι της οδηγούν σε σελίδες με αξία περιεχομένου. Μετά από κάθε κύκλο γίνεται ομαλοποίηση των τιμών διαιρώντας το σκορ του κόμβου με το τετράγωνο του συνόλου των σκορ όλων των κόμβων, μια για το περιεχόμενο και μια για τις συνδέσεις.

## **Υπολογισμός πυκνότητας λέξεων-κλειδιών**

Για τον υπολογισμό της πυκνότητας των λέξεων-κλειδιών υπάρχουν πολλά επαγγελματικά εργαλεία στο διαδίκτυο, άλλα δωρεάν και άλλα επί πληρωμής, όπως αναφέρθηκαν στο 4<sup>ο</sup> κεφάλαιο της πτυχιακής. Για όσους χρήστες δεν είναι τόσο εξοικειωμένοι ή προτιμούν ένα πιο γρήγορο τρόπο μέτρησης της πυκνότητας μιας σελίδας, μπορούν να χρησιμοποιήσουν το Microsoft Word.

Για τον έλεγχο της πυκνότητας των λέξεων-κλειδιών με το Word ακολουθείται η εξής διαδικασία:

1. Αντιγραφή ολόκληρου του κειμένου της σελίδας στο Microsoft Word.
2. Από το μενού 'Εργαλεία' επιλέγετε την 'Καταμέτρηση Λέξεων'. Σημειώστε τον συνολικό αριθμό των λέξεων στην σελίδα.
3. Από το μενού 'Επεξεργασία' επιλέγετε 'Αντικατάσταση'. Συμπληρώστε στο πεδίο 'Εύρεση του' την λέξη-κλειδί και στο πεδίο 'Αντικατάσταση με' την λέξη-κλειδί (ώστε να μην αλλάξει το περιεχόμενο της σελίδας). Πα-

τήστε το πλήκτρο 'Αντικατάσταση όλων' και σημειώστε τον αριθμό των αντικαταστάσεων που έγιναν.

4. Διαιρέστε τον αριθμό των αντικαταστάσεων με τον αριθμό των λέξεων που είχατε βρει νωρίτερα και έχετε υπολογίσει την πυκνότητα των λέξεων-κλειδιών.