



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

Ημι-αυτόματη Γεωκωδικοποίηση Ιστοσελίδων

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

της

ΧΑΡΙΚΛΕΙΑΣ ΛΟΝΤΟΥ

Επιβλέπων : Τίμος Σελλής
Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2008



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

Ημι-αυτόματη Γεωκωδικοποίηση Ιστοσελίδων

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

της

ΧΑΡΙΚΛΕΙΑΣ ΛΟΝΤΟΥ

Επιβλέπων : Τίμος Σελλής
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 9^η Ιουλίου 2008

.....
Τίμος Σελλής
Καθηγητής Ε.Μ.Π.

.....
Ιωάννης Βασιλείου
Καθηγητής Ε.Μ.Π.

.....
Νεκτάριο Κοζύρη
Αναπλ. Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2008

.....

ΛΟΝΤΟΥ ΧΑΡΙΚΛΕΙΑ

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Χαρίκλεια Δ. Λόντου, 2008.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ' ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Η παρούσα διπλωματική εργασία αναπτύσσει την μεθοδολογία της ημι-αυτόματης γεωκωδικοποίησης ιστοσελίδων που μακροπρόθεσμα το περιεχόμενό τους δεν αλλάζει, δηλαδή την απόδοση γεωγραφικών συντεταγμένων στις γεωγραφικές οντότητες που υπάρχουν στο κείμενο τους χρησιμοποιώντας έναν συνδυασμό αυτόματης και βασιζόμενη στον άνθρωπο γεωκωδικοποίησης.

Οι αλγόριθμοι που χρησιμοποιούνται για το geoparsing και την γεωκωδικοποίηση έχουν εξαιρετικά αποτελέσματα όταν εφαρμόζονται για τον προσδιορισμό αριθμών τηλεφώνων και διευθύνσεων, όμως σε πιο γενικευμένα δεδομένα, οι αλγόριθμοι αυτοί εμφανίζουν ένα σημαντικό ποσοστό λάθους τόσο στο να βρίσκουν οντότητες που δεν θα έπρεπε (Ελ. Βενιζέλος είναι πρόσωπο), όσο και να μην βρίσκουν άλλες που θα έπρεπε (Ελ. Βενιζέλος- το διεθνές αεροδρόμιο της Αθήνας).

Η παρούσα διπλωματική εργασία χρησιμοποιεί την ανθρώπινη παρέμβαση ώστε να βελτιώσει τα αποτελέσματα της αυτόματης γεωκωδικοποίησης. Για τον σκοπό αυτό αναπτύχθηκε ένα Web browser extension το οποίο α) επιτρέπει την χειρωνακτική γεωκωδικοποίηση λέξεων ή φράσεων και β) την βελτίωση των αποτελεσμάτων, συμπεριλαμβανομένης και της διαγραφής των λανθασμένων αποτελεσμάτων. Αυτή η προτεινόμενη προσέγγιση είναι ιδιαίτερα χρήσιμη για ιστοσελίδες όπως η Βικιπαιδεία, δηλ., σελίδες που έχουν μια ορισμένη αξία στην κοινότητα, ιδιαίτερα προσεγμένες και που το περιεχόμενό τους αλλάζει σχετικά αργά. Εδώ, η γεωκωδικοποίηση μπορεί να αποτελέσει μέρος της δημιουργίας μιας ιστοσελίδας.

Τα αποτελέσματα από την γεωκωδικοποίηση αποθηκεύονται σε μια βάση δεδομένων, δηλ., αποθηκεύονται το URL της ιστοσελίδας, η θέση της λέξης ή της φράσης και οι αντίστοιχες συντεταγμένες και τέλος η μέρα της γεωκωδικοποίησης, δηλ., η έκδοση της ιστοσελίδας. Τα αποτελέσματα απεικονίζονται υπογραμμίζοντας τις αντίστοιχες λέξεις ή φράσεις και αντιστοιχίζοντας τους χάρτες που τις περιγράφουν. Στην περίπτωση μας χρησιμοποιήθηκαν οι χάρτες της Google. Στο τεχνικό μέρος, χρησιμοποιήθηκαν ένα αυτόματο εργαλείο γεωκωδικοποίησης, που αναπτύχθηκε από τον σπουδαστή Αλβέρτο-Δαβίδ, Α. Άντζελ, ο Apache Web server καθώς και ένα browser extension σε γλώσσα JavaScript και τέλος η γλώσσα java για την θεμελίωση της επικοινωνίας.

Λέξεις Κλειδιά: Χειρωνακτική γεωκωδικοποίηση, ημι-αυτόματη γεωκωδικοποίηση, απεικόνιση γεωγραφικής πληροφορίας, extension

Abstract

This thesis presents a methodology for the semi-automatic geocoding of persistent Web pages, i.e., relating identifiers in texts to geographic co-ordinates using a combined automatic and human-centered approach.

Geoparsing and geocoding algorithms are successfully applied to identify phone numbers and addresses, however when more generic geo identifiers are involved, automatic algorithms produce a significant number of false positives (Venizelos as a person) and false negatives (Venizelos as the name of Athens international airport).

This thesis advocates human intervention to improve on automatic geocoding results and develops therefore a Web browser extension that (i) allows for the manual geocoding of text portions and (ii) the updating, including deletion of automatically generated results. This proposed approach is especially helpful for persistent Web pages such as Wikipedia, i.e., pages that have a certain value to the community, are well cared for and change rather slowly. Here, geocoding can become a regular part of Web page authoring.

The geocoding of a Web page is stored in a database, i.e., a Web page is stored in terms of its URL, the geocoded text portions in terms of their position on the Web page and the respective co-ordinates, and the date of the geocoding, i.e., the version of the Web page. Further, the geocoding of a Web page is displayed as highlighted text and by means of a map, i.e., clicking on a text portions shows the respective position on a map. In our case, Google Maps was used for this task. In terms of technology, the developed system includes the automatic geocoding tool developed by Albert-David Angel, an Apache Web server and Tomcat servlet engine, as well as the browser extension developed in java script and java for the interaction with the servlet. This work also includes a set of use cases that illustrate the applicability and usefulness of the overall approach.

Keywords: manual geocoding, semi-automatic geocoding, depiction of geographic information, extension.

Ευχαριστίες

Αρχικά θα ήθελα να ευχαριστήσω όλους όσους συνέβαλαν για την ολοκλήρωση αυτού του τόμου, με πρώτο και κύριο τον επιβλέποντα της διπλωματικής μου Dieter Pfoser, που διαίθεσε πολύτιμο χρόνο για να με κατευθύνει, να με εμπνεύσει μα κυρίως για να ακούσει ότιδήποτε έλεγα, όσο κουτό και αν ήταν, με μεγάλη προσοχή.

Θα ήθελα επίσης να ευχαριστήσω τον καθηγητή κ. Τιμο Σελλή που ενέκρινε την πορεία της διπλωματικής μου εργασίας και που αποτελεί πρότυπο με την στάση του απέναντι σε όλους του φοιτητές .

Η αλήθεια είναι όμως πως χωρίς την στήριξη των γονίων μου τίποτα από όλα αυτά δεν θα είχε συμβεί. Τέλος, ένα μεγάλο ευχαρίστω στους φίλους και όσους με βοήθησαν και με εμπνεύχωναν, ο κάθε ένας με τον δικό του μοναδικό τρόπο.

Πίνακας περιεχομένων

1	Εισαγωγή.....	1
1.1	Γεωκωδικοποίηση Ιστοσελίδων.....	1
1.2	Οργάνωση τόμου.....	2
2	Σχετικές εργασίες.....	3
2.1	Βασικές έννοιες	3
2.1.1	<i>Geoparsing</i>	3
2.1.2	<i>Geocoding</i>	4
2.2	Δημιουργία μιας καινούργιας ανάγκης	4
2.2.1	<i>Διάφορα προβλήματα</i>	5
2.3	Άλλες σχετικές εργασίες	6
2.3.1	<i>Εμπορικές προτάσεις</i>	6
2.3.2	<i>Ερευνητικές Προτάσεις</i>	9
2.4	Σκοπός της διπλωματικής.....	11
3	Αυτόματη Γεωκωδικοποίηση	13
3.1	Σκοπός της διπλωματικής.....	13
3.2	Αρχιτεκτονική Συστήματος.....	14
3.2.1	<i>Geoparsing</i>	15
3.2.2	<i>Γεωκωδικοποίηση - Geocoding</i>	15
3.2.2.1	Παράδειγμα γεωκωδικοποίησης.....	15
3.2.2.2	Ολοκλήρωση αποτελεσμάτων.....	16
3.2.3	<i>Γεοδεδομένα - GeoData</i>	16
3.3	Παραδείγματα.....	16
3.3.1	<i>Παράδειγμα «Κόρινθος»</i>	16
3.3.2	<i>Παράδειγμα «Λεόντειο Αχαΐας»</i>	19
3.4	Προβλήματα.....	21
3.4.1	<i>Το πρόβλημα της ταυτότητας του αντικείμενου</i>	21
3.4.1.1	Αιτίες του προβλήματος ταυτότητας αντικείμενου.....	22
3.4.2	<i>Προσεγγιστικό ταίριασμα συμβολοσειρών</i>	22

3.4.2.1 Το πρόβλημα των Greeklish.....	22
3.4.3 Προσεγγιστική αναζήτηση συμβολοσειρών.....	23
3.4.4 Το Πρόβλημα των Γεωδεδομένων.....	23
3.4.4.1 Έλλειψη δεδομένων.....	23
3.4.4.2 Ποιότητα δεδομένων.....	24
3.4.4.3 Καθαρισμός δεδομένων.....	24
3.5 Η δική μας μετατροπή	25
4 Ανάλυση Απαιτήσεων Συστήματος.....	27
4.1 Αρχιτεκτονική	27
4.2 Περιγραφή Λειτουργιών.....	29
4.2.1 <i>Browser Extension (client side)</i>	29
4.2.2 <i>Server (server side)</i>	30
4.3 Ανάλυση Επιμέρους Τμημάτων.....	31
4.3.1 User Interface - Browser Extension	32
4.3.2 Java Servlet.....	34
4.3.3 Visualization - Google Maps.....	35
4.3.4 Persistent Storage - Βάση Δεδομένων.....	36
5 Λεπτομέρειες Υλοποίησης	37
5.1 <i>Mozilla FireFox</i>	37
5.2 <i>Mozilla extension</i>	37
5.2.1 Το Hello, World! του extension.....	38
5.2.1.1 Σχεδίαση του Extension	39
5.2.1.2 Επεξήγηση Αρχείων.....	40
5.2.1.3 Αναγνώριση του Extension από τον Διαχειριστή των Extensions.....	44
5.2.1.4 «Πακετάρισμα».....	45
5.2.2 Το δικό μας Extension.....	46
5.3 <i>Google MAPS API</i>	48
5.3.1 Προσφερόμενες υπηρεσίες.....	49
5.3.2 <i>To “Hello World” στο Google Maps</i>	52
5.3.2.1 Loading το Google Maps API.....	53
5.3.2.2 Τα στοιχεία DOM του χάρτη.....	53
5.3.2.3 GMap2 – Το κύριο αντικείμενο.....	53
5.3.2.4 Αρχικοποίηση του χάρτη.....	54
5.3.2.5 Loading the Map.....	54

5.3.2.6 Latitudes και Longitudes	54
5.3.2.7 Παράθυρα Πληροφοριών – Info Window.....	54
5.3.3 Οι Google maps της εφαρμογής μας.....	55
<i>Κώδικας για την Προσθήκη Γεωγραφικής Πληροφορίας.....</i>	<i>57</i>
5.4 JavaScript.....	57
5.5 Postgres SQL.....	58
5.6 Tomcat.....	58
5.7 Servlets	59
5.8 Java.....	61
5.9 AJAX.....	61
5.10 PHP.....	62
6 Παρουσίαση Αποτελεσμάτων.....	63
6.1 Οδηγός Χρήσης	63
6.1.1 Απεικόνιση Γεωγραφικής πληροφορίας.....	63
6.1.2 Διόρθωση Λαθών	65
6.1.2.1 Αλλαγή Συντεταγμένων	65
6.1.2.2 Διαγραφή Γεωγραφικής Οντότητας.....	66
6.1.2.3 Προσθήκη Νέας Γεωγραφικής Οντότητας.....	68
6.2 Μερικά Παραδείγματα.....	70
7 Επίλογος	75
7.1 Σύνοψη και συμπεράσματα.....	75
7.2 Μελλοντικές επεκτάσεις.....	75
7.2.1 Δημιουργία μίας δημόσιας αποθήκης ή υπηρεσίας για την γεωκωδικοποίηση ιστοσελίδων	76
7.2.2 Λογισμικό Ανοιχτού Κώδικά.....	76
8 Βιβλιογραφία.....	77

1

Εισαγωγή

1.1 Γεωκωδικοποίηση Ιστοσελίδων

Τα τελευταία χρόνια έχει γίνει μεγάλη προσπάθεια στην δημιουργία εργαλείων που θα μπορούν να εξάγουν πληροφορίες από το περιεχόμενο των ιστοσελίδων (semantic web). Η προσπάθεια αυτή εστιάζεται κυρίως στις γεωγραφικές πληροφορίες που ιστοσελίδες μπορούν να προσφέρουν, παρέχοντας στον χρήστη υπηρεσίες όπως αναζήτησης ιστοσελίδων με βάση γεωγραφικά κριτήρια. Έτσι, λοιπόν, οι ιστοσελίδες μπορούν να καταταχθούν με βάση τις γεωγραφικές πληροφορίες που βρίσκονται στο περιεχόμενό τους.

Όλα τα εργαλεία που χρησιμοποιούνται μέχρι σήμερα για τον παραπάνω σκοπό στηρίζονται εξολοκλήρου στην αυτόματη εξαγωγή των γεωγραφικών πληροφοριών από μια ιστοσελίδα, κάτι που δημιουργεί προβλήματα κυρίως σε γλωσσολογικό επίπεδο, αφού μέχρι και σήμερα καμία μηχανή δεν μπορεί να κατανοήσει πλήρως ένα κείμενο, παρά τις αξιόλογες προσπάθειες που έχουν λάβει χώρα. Οι αλγόριθμοι που χρησιμοποιούνται έχουν εξαιρετικά αποτελέσματα όταν εφαρμόζονται για τον προσδιορισμό αριθμών τηλεφώνων και διευθύνσεων, όμως σε πιο γενικευμένα δεδομένα, οι αλγόριθμοι αυτοί εμφανίζουν ένα σημαντικό ποσοστό λάθους τόσο στο να βρίσκουν οντότητες που δεν θα έπρεπε (Ελ. Βενιζέλος είναι πρόσωπο), όσο και να μην βρίσκουν άλλες που θα έπρεπε (Ελ. Βενιζέλος- το διεθνές αεροδρόμιο της Αθήνας).

Ορμώμενη από τα προηγούμενα στην παρούσα διπλωματική εργασία δημιουργήσαμε ένα εργαλείο που επιτρέπει την ανθρώπινη παρέμβαση ώστε να βελτιωθούν τα αποτελέσματα της

αυτόματης γεωκωδικοποίησης. Για τον σκοπό αυτό αναπτύχθηκε ένα Web browser extension το οποίο α) επιτρέπει την χειρωνακτική γεωκωδικοποίηση λέξεων ή φράσεων και β) την βελτίωση των αποτελεσμάτων, συμπεριλαμβανομένης και της διαγραφής των λανθασμένων αποτελεσμάτων. Όπως, είναι εύκολο να διαπιστώσει κανείς η εφαρμογή αυτή είναι ιδιαίτερα χρήσιμη κυρίως σε ιστοσελίδες που το περιεχόμενό τους αλλάζει με αργό ρυθμό ή μεταβάλλεται ελάχιστα όπως η Βικιπαιδεία.

1.2 Οργάνωση τόμου

Η διπλωματική εργασία αποτελείται από 8 κεφάλαια, στα οποία γίνεται ανάλυση του θεωρητικού υπόβαθρου της παρούσας εργασίας, των εργαλείων και των μεθόδων που χρησιμοποιήθηκαν.

Το κεφάλαιο 2 ασχολείται με την περιγραφή των αιτίων που οδήγησαν στην ανάγκη για την γεωκωδικοποίηση του παγκόσμιου Ιστού και των κύριων προβλημάτων που πρέπει να αντιμετωπιστούν. Επίσης, γίνεται μια συνοπτική αναφορά των σημαντικότερων εμπορικών και ερευνητικών προσπαθειών στον τομέα της γεωκωδικοποίησης. Και κλείνει με μια την αποσαφήνιση το κύριο σκοπού της εργασίας.

Στο κεφάλαιο 3 περιγράφεται το εργαλείο αυτόματης γεωκωδικοποίησης ιστοσελίδων που κατασκευάστηκε από τον σπουδαστή Αλβέρτο-Δαυΐδ, Α. Άντζελ για την διπλωματική εργασία “Εξαγωγή Γεωγραφικής Πληροφορίας από Ημιδομημένο Κείμενο”. Επισημαίνονται τα βασικότερα προβλήματα του εργαλείου αυτού, προβλήματα που οδήγησαν στην δημιουργία της παρούσας Διπλωματικής εργασίας.

Στο κεφάλαιο 4 γίνεται περιγραφή της λογικής πάνω στην οποία είναι χτισμένο το σύστημα. Αναλύεται η αρχιτεκτονική του και διαχωρίζονται και επεξηγούνται τα υποσυστήματα που το αποτελούν.

Στο 5ο κεφάλαιο βλέπουμε τα στοιχεία της υλοποίησης, δηλαδή το πως δουλεύει σε προγραμματιστικό επίπεδο το κάθε τμήματα του συστήματος και αναλύονται οι πλατφόρμες και οι γλώσσες προγραμματισμού που χρησιμοποιήθηκαν.

Το κεφάλαιο 6 αποτελείται από το δύο ξεχωριστά τμήματα. Το πρώτο είναι ένα αναλυτικό εγχειρίδιο χρήστη του εργαλείου μας και το δεύτερο περιλαμβάνει κάποια παραδείγματα χρήσης του καθώς και παραδείγματα σε ιστοσελίδες όπου τα αποτελέσματα έχουν επεξεργαστεί από τον χρήστη, συγκρίνοντας με αυτά του αυτόματου εργαλείου του κ. Άντζελ.

Το 7ο κεφάλαιο αποτελεί τον επίλογο της εργασίας, συνθέτοντας τα συμπεράσματα αυτής, και δίνοντας μερικές κατευθύνσεις για μελλοντικές επεκτάσεις.

Κλείνοντας, στο κεφάλαιο 8 παρουσιάζεται ενδεικτική βιβλιογραφία για όλα όσα εξετάζονται στον τόμο αυτό.

2

Σχετικές εργασίες

Το κεφάλαιο αυτό περιγράφει τους λόγους που οδήγησαν στην ανάγκη για γεωκωδικοποίηση των ιστοσελίδων και αποτελεί μια ανασκόπηση των σημαντικότερων εφαρμογών του χώρου. Τέλος, γίνεται μία αναφορά στον σκοπό της παρούσας διπλωματικής εργασίας και στον σκοπό στον οποίο αποσκοπεί.

2.1 Βασικές έννοιες

Κεντρικές έννοιες στην εργασία αυτή αποτελούν το *geoparsing*, και η *γεωκωδικοποίηση* (geocoding). Στην συνέχεια ακολουθεί σύντομη περιγραφή τους, έτσι ώστε ο αναγνώστης να κατατοπιστεί με το τι ακριβώς πραγματεύονται αυτοί οι όροι

2.1.1 Geoparsing

Το *geoparsing* σημαίνει λεκτική και γραμματική ανάλυση μιας ιστοσελίδας, πιθανώς υποβοηθούμενη από κάποιο *gazetteer* (ευρετήριο τοπωνυμίων), με στόχο την ανακάλυψη πιθανής γεωγραφικής πληροφορίας. Η πληροφορία αυτή μπορεί να εμφανίζεται με διάφορες μορφές (από το επίπεδο του δικτύου, όπως διεύθυνση IP, από επίπεδο κειμένου, όπως τηλέφωνα, ταχυδρομικοί κώδικες, από επίπεδο σημασιολογίας, όπως διευθύνσεις, τοπωνύμια, κατευθύνσεις, από επίπεδο τοπολογίας κ.α.).

2.1.2 Geocoding

Η *γεωκωδικοποίηση* είναι η μετατροπή αυτής της πληροφορίας σε συντεταγμένες, σε κάποιο γεωγραφικό σύστημα αναφοράς. Αν και αυθύπαρκτη διαδικασία, στα πλαίσια της εργασίας αυτής θεωρήθηκε ως φυσική συνέχεια του *geoparsing*.

2.2 Δημιουργία μιας καινούργιας ανάγκης

Από τις πρώτες μέρες του παγκόσμιου Ιστού, πολλές προσπάθειες έχουν γίνει για την οργάνωση, την πλοήγηση, και την αναζήτηση των εγγράφων που φιλοξενούνται από αυτόν. Οι γνώστες μας αναζητήσεις με την βοήθεια μιας “λέξης-κλειδί”, παρόλο που δεν αποτελούν πλέον μια και τόσο απλή διαδικασία αφού ο Ιστός έχει γίνει μεγαλώσει υπερβολικά, συνεχίζουν να παρέχουν μια από τις πιο χρήσιμες τεχνικές για τους χρήστες για να βρουν πληροφορίες για ένα δεδομένο θέμα, αλλά επίσης αποτελούν την αφετηρία για διάφορες υψηλότερου επιπέδου σημασιολογικές ερωτήσεις που μπορούν να χρησιμοποιηθούν σε μια αναζήτηση πληροφοριών. Για παράδειγμα, ένας αναγνώστης μπορεί ενδιαφέρεται μόνο για την εύρεση εγγράφων γραμμένα στα ιαπωνικά, ή για έγγραφα που δημιουργήθηκαν ή τροποποιήθηκαν μετά από μια ορισμένη ημερομηνία. Τέτοιοι στόχοι ανάκτησης πληροφοριών δεν εξυπηρετούνται από μια απλή αναζήτηση “λέξης-κλειδί”, και τα συστήματα που οργανώνονται για να απαντήσουν σε τέτοιες αναζητήσεις επιθυμούν συνήθως υψηλότερου επιπέδου σημασιολογικές πληροφορίες από το περιεχόμενο των ιστοσελίδων.

Είναι γεγονός ότι κάθε ιστοσελίδα περιέχει άμεσα ή έμμεσα κάποια γεωγραφικά στοιχεία. Έτσι οι χρήστες θα μπορούσαν να προσθέσουν *γεωγραφικά κριτήρια* στις αναζητήσεις τους βοηθώντας στον πιο εύκολο εντοπισμό ιστοσελίδων που τους αφορούν. Παραδείγματα τέτοιων χρηστών αποτελούν οι ταξιδιώτες που επιθυμούν να μάθουν τι ευκαιρίες υπάρχουν σε έναν προορισμό, οι σπουδαστές που μελετούν για ένα μέρος του κόσμου, αναλυτές που ετοιμάζουν μια έκθεση σχετικά με μια δεδομένη περιοχή, Επιχειρήσεις που ερευνούν μια περιοχή για επέκταση των εγκαταστάσεων τους, ή κυβερνητικοί αρμόδιοι που μελετούν τον περιβαλλοντικό ή οικονομικό αντίκτυπο σε μια συγκεκριμένη περιοχή.

Επίσης η *γεωγραφική διανομή* των ιστοσελίδων θα μπορούσε να βρεθεί, για παράδειγμα, επιδεικνύοντας πως ένα ορισμένο προϊόν είναι δημοφιλές σε μια ορισμένη γεωγραφική περιοχή και μη δημοφιλές σε άλλες. Ο συσχετισμός μεταξύ των αναφορών σε διάφορες τοποθεσίες, ή οι ίδιες οι τοποθεσίες, θα μπορούσαν να αναλυθούν, παραδείγματος χάριν, για να βρεθούν ποιες περιοχές είναι συνδεδεμένες από τους bloggers με τη μόδα, τα πάρτι, τις διακοπές, ή το καλό φαγητό. Μια ποικιλία από υπηρεσίες που βασίζονται στην τοποθεσία μπορούν πλέον να αναπτυχθούν για τις κινητές συσκευές.

Για να επιτευχθούν όλοι οι παραπάνω στόχοι, πρέπει να βρεθεί ο γεωγραφικός προσανατολισμός της ιστοσελίδας ή το γεωγραφικό πλαίσιο μέσα στο οποίο αυτή κινείται. Η πιο ενδιαφέρουσα και σημαντική πηγή για το γεωγραφικό πλαίσιο μιας ιστοσελίδας παρέχεται από το ίδιο το *περιεχόμενο* της. Παραδείγματος χάριν, η ιστοσελίδα μιας επιχείρησης μπορεί να περιέχει τη διεύθυνση ή τον τηλεφωνικό αριθμό της έδρας της, που αποτελεί γεωγραφικό στοιχείο μέσα στη σελίδα. Σε μια σελίδα ειδήσεων πρέπει να υπάρχει ένα γεωγραφικό στοιχείο της θέσης όπου το γεγονός πραγματοποιήθηκε. Όλα αυτά αποτελούν ισχυρά στοιχεία για την εύρεση του γεωγραφικού πλαισίου της ιστοσελίδας. Επίσης, ιδιαίτερα σημαντικό ρόλο στην διαμόρφωση της ιστοσελίδας παίζει και η προέλευση της, η φυσική θέση του κεντρικού υπολογιστή που αποθηκεύεται, η διεύθυνση του συντάκτη ή του ιδιοκτήτη της, κ.λ.π.

2.2.1 Διάφορα προβλήματα

Πολλές ιστοσελίδες όμως δεν έχουν κάποιο ισχυρό γεωγραφικό στοιχείο, έτσι θα ήταν παραπλανητικό να περιμένουμε τέτοια στοιχεία για κάθε ιστοσελίδα. Η ισχύς ενός γεωγραφικού στοιχείου είναι δύσκολο να καθοριστεί, αλλά πολλαπλοί δείκτες μπορούν να χρησιμοποιηθούν για την βελτίωση των αποτελεσμάτων. Επιπλέον, μερικοί δείκτες δίνουν υψηλότερη ισχύς από άλλους. Για παράδειγμα, συγκρίνοντας τη φάση «*Κασπία θάλασσα*» μέσα σε μία ιστοσελίδα με έναν τηλεφωνικό αριθμό που βρίσκεται στην ακτή της Κασπίας θάλασσας. Ο τηλεφωνικός αριθμός είναι ένας τρόπος επικοινωνίας με ένα πρόσωπο που έχει τις συμπληρωματικές πληροφορίες για το πραγματικό θέμα της σελίδας, αλλά στην πραγματικότητα να μην έχει καμία σχέση με την περιοχή. Σε αντίθεση, στην φράση «*Κασπίας θάλασσα*», εκτός αν έχει μια άλλη αναγνωρίσιμη έννοια, η αμφιβολία για το αν αποτελεί γεωγραφικό στοιχείο είναι ελάχιστη. Ένας γεωγραφικός δείκτης μπορεί να τυχαίος, όμως ακόμα και έτσι παρέχει έναν τρόπο να ταξινομηθούν τα έγγραφα σύμφωνα με τη γεωγραφική σχετικότητά τους. Εκείνα που περιέχουν έναν ισχυρό γεωγραφικό δείκτη οι πληροφορίες που περιέχουν είναι πιο σωστές από άλλα που δεν έχουν.

Το πιο σημαντικό πρόβλημα που προκύπτει είναι ίσως η αποσαφήνιση των γεωγραφικών ονομάτων. Διακρίνουμε δύο περιπτώσεις: την *geo/non-geo* και τη *geo/geo*. Μια ασάφεια *geo/non-geo* εμφανίζεται όταν το όνομα μιας τοποθεσίας έχει και μη-γεωγραφική έννοια, όπως ένα όνομα προσώπων (π.χ. Ελ. Βενιζέλος) ή μια κοινή λέξη (Αετός-όνομα δήμου στο νομό Μεσσηνίας και όνομα πουλιού). Στο *geo/geo* η ασάφεια προκύπτει όταν περισσότερες από μια περιοχές έχουν το ίδιο όνομα, όπως Λονδίνο, Αγγλία και Λονδίνο, Οντάριο. Το πρόβλημα της γεωγραφικής αποσαφήνισης ονόματος έχει αντιμετωπιστεί σχεδόν ανεξάρτητα από τους κύκλους της γλωσσολογίας, της κατανόησης μηχανών και της τεχνητής νοημοσύνης, καθώς επίσης και από τους ερευνητές στον τομέα της ανάκτησης πληροφοριών και της διαχείρισης γνώσης, με αποτέλεσμα τις διαφορετικές λύσεις. Ενώ σύμφωνα με την

ΕΦ¹ η μηχανή αναγνωρίζει τα ονόματα από τη δομή και το περιεχόμενο του κειμένου στο οποίο βρίσκονται, η προσέγγιση της ανάλυσης δεδομένων προσφεύγει σε γλωσσάρια και γεωγραφικά λεξικά για να τα προσδιορίσει, και να πει σε ποια θέση αναφέρονται. Η τελευταία προσέγγιση δεν μπορεί να βρει τα ονόματα που δεν είναι στον κατάλογο, όπως η προηγούμενη μπορεί, αλλά οδηγεί σε έναν απλούστερο αλγόριθμο. Επίσης δεν απαιτεί *training data* που είναι μερικές φορές δύσκολο να έρθει κοντά. Επιπλέον, δεδομένου ότι η ανάλυση δεδομένων από τον Ιστό συνεπάγεται τεράστιες ποσότητες στοιχείων, ο αλγόριθμος που θα χρησιμοποιηθεί πρέπει να είναι αρκετά γρήγορος.

Είναι εύλογο να αναρωτηθεί κανείς πιο ποσοστό ιστοσελίδων περιέχει ένα αναγνωρίσιμο γεωγραφικό στοιχείο. Με βάση τα πειράματα με έναν αρκετά μεγάλο δείγμα ιστοσελίδων που πραγματοποιήθηκαν για το 8, βρέθηκε ότι περίπου **4.5%** των συνολικών ιστοσελίδων περιέχουν έναν αναγνωρίσιμο US zip code, **8.5%** περιέχουν έναν αναγνωρίσιμο τηλεφωνικό αριθμό, και **9.5%** περιέχουν τουλάχιστον ένα από τα δύο. Τα πραγματικά ποσοστά μπορεί να είναι κάπως μεγαλύτερα – αφού αυτά υπολογίστηκαν με βάση τον parser που δημιουργήθηκε για τις ανάγκες του 8.

Τα εργαλεία πλοήγησης μπορούν να στηριχθούν στις πληροφορίες που έχουν εξαχθεί κατά τη διάρκεια επεξεργασίας των ιστοσελίδων και των αποτελεσμάτων που έχουν αποθηκευτεί σε μια βάση δεδομένων.

2.3 Άλλες σχετικές εργασίες

Προκειμένου να οριοθετήσουμε τον χώρο που πραγματεύεται η εργασία αυτή, χρήσιμη θα ήταν μια σύντομη ανασκόπηση ορισμένων αποτελεσμάτων που παρουσιάζουν ενδιαφέρον. Θα δούμε, συνοπτικά, κάποιες ολοκληρωμένες προτάσεις, ερευνητικές και εμπορικές.

2.3.1 Εμπορικές προτάσεις

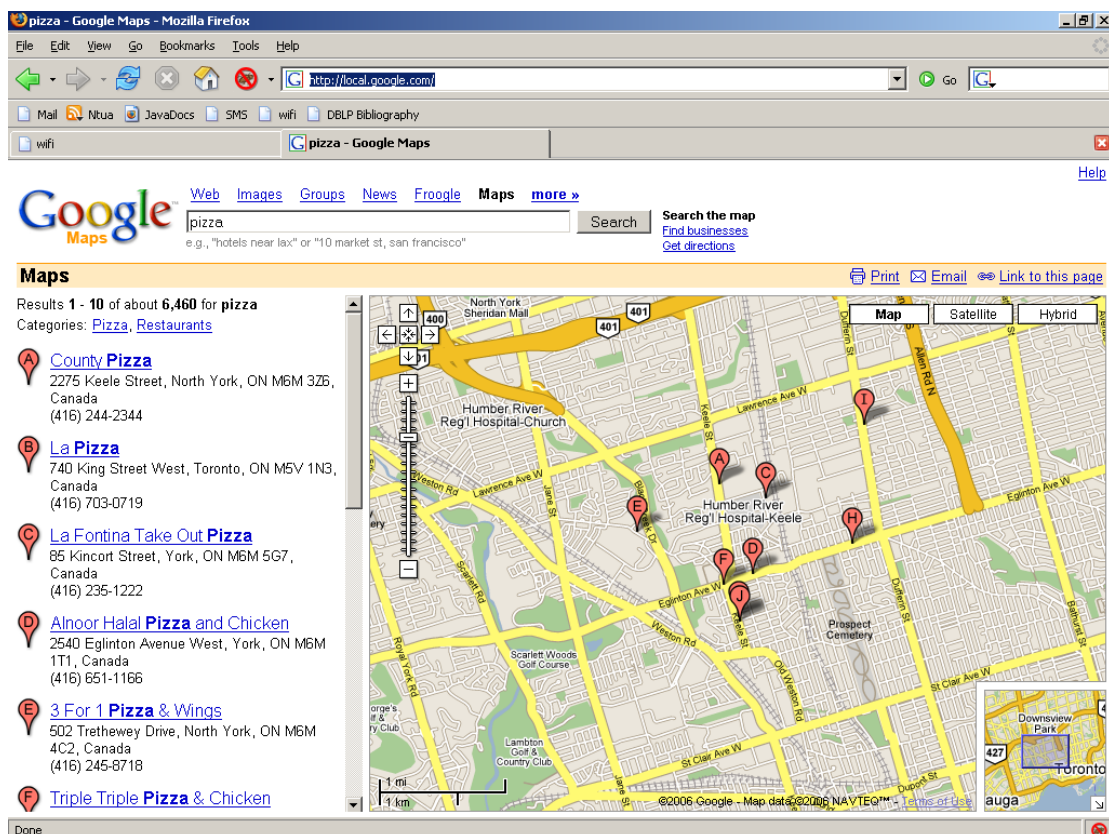
Οποιαδήποτε επισκόπηση εφαρμογών "εξαγωγής πληροφορίας" θα είναι ελλιπής χωρίς την αναφορά του κολοσσού που ακούει στο όνομα Google^{TM2}. Τον τελευταίο χρόνο, το Google έχτισε, και έκτοτε αναβαθμίζει διαρκώς, την υπηρεσία Google Local8 (περιλαμβάνει και το Google Maps⁸κάτω και Google Earth⁸). Με την υπηρεσία αυτή ο δ χρήστης μπορεί, μεταξύ άλλων, να αναζητήσει με βάση γεωγραφικά κριτήρια σε συνδυασμό με λέξεις-κλειδιά ("βρες πιτσαρίες σε αυτήν την περιοχή του χάρτη"). Η δυνατότητα αυτή, βέβαια, λειτουργεί

¹ ή νευρο-γλωσσολογικός προγραμματισμός, είναι η μελέτη για το πώς οι άνθρωποι αντιπροσωπεύουν την εμπειρία, πραγματικός ή φαντασμένος, σε τους εσωτερικά, και την αντίστοιχη επίδραση στον τρόπο με τον οποίο το νευρικό σύστημα λειτουργεί.

² Στο 8 αναφέρεται, σχεδόν ανεκδοτολογικά, ότι "Μια από τις καλύτερες ερωτήσεις που μπορείτε να ρωτήσετε (τον εαυτό σας), όσο σκέφτεστε τη διαδικασία της ανάλυσης, είναι "Τι θα έκανε το Google;". Αν και οι αλγόριθμοι του Google είναι κλειστοί και κρατιούνται αρκετά μυστικοί, τα αποτελέσματα των αναζητήσεων βοηθούν στην κατανόησή του."

ικανοποιητικά, προς το παρόν, μόνο για τις Ηνωμένες Πολιτείες. Εξετάζοντας την λειτουργία της υπηρεσίας, παρατηρούμε ότι λειτουργεί ψάχνοντας στις σελίδες που, ούτως ή άλλως η Google έχει στα ευρετήριά της, ταχυδρομικούς κώδικες και τοπωνύμια, ενώ κάνει και χρήση κατηγοριοποιήσεων των ιστοσελίδων από ανθρώπους³.

Αντίστοιχα, ίσως με λιγότερο εντυπωσιακό παρουσιαστικό, φαίνεται να λειτουργεί ο



Σχήμα 2.1 : Ψάχνοντας για πίτσα στο Toronto με το Google Local 8

χρυσός οδηγός του Yahoo™, αν και εδώ οι πληροφορίες προέρχονται μονάχα από την ανθρώπινη κατηγοριοποίηση ιστοσελίδων.

Ένα άλλο εμπορικό προϊόν που εντοπίζει έναν εξυπηρετητή ιστού στο χάρτη (με βάση τη διεύθυνση IP του, την καταχώρησή του στην υπηρεσία WhoIs, κ.α.), είναι το NeoTrace8.

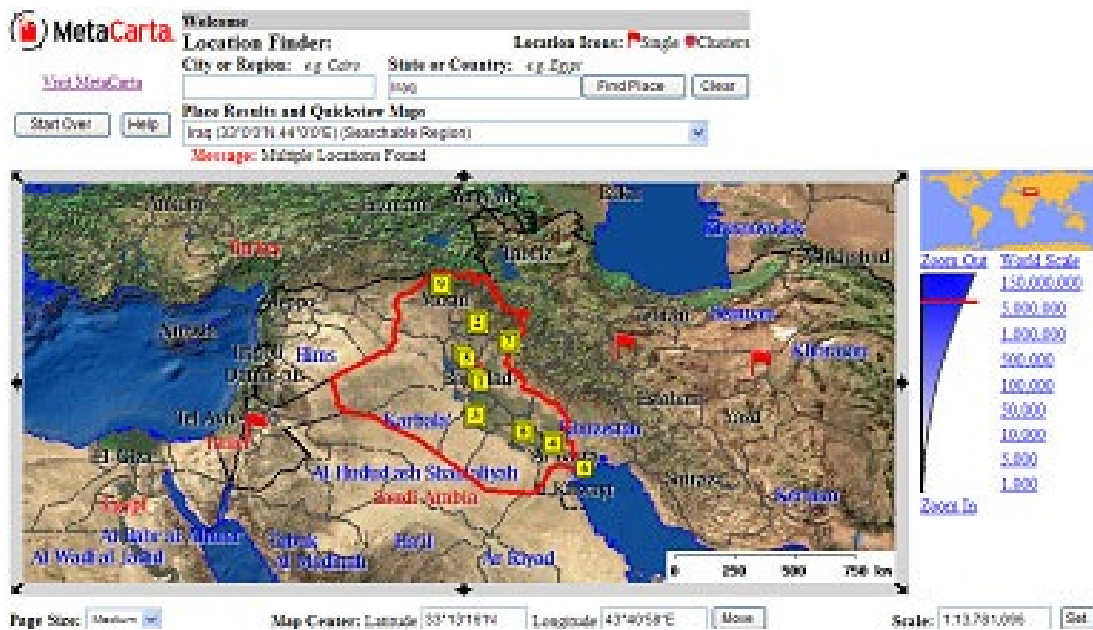
Περισσότερο ενδιαφέρον, για εμάς, παρουσιάζουν οι, εμπορικές, υπηρεσίες της MetaCarta8. Σε αυτές βρίσκουμε μια ενοποιημένη αντιμετώπιση geoparsing/γεωκωδικοποίησης, που χρησιμοποιεί στοιχεία επεξεργασίας φυσικής γλώσσας, στατιστικής, και γεω-δεδομένα υψηλής λεπτομέρειας. Τα αποτελέσματα μπορούν να χρησιμοποιηθούν για γεωγραφικές αναζητήσεις στο διαδίκτυο, σε εφαρμογές GIS, για γεωγραφική ομαδοποίηση/συσταδοποίηση εγγράφων κ.ο.κ. Σημαντική είναι επίσης η προσφορά της με το

³ Δηλαδή, κάποιος υπάλληλος της εταιρίας πρέπει να εξετάσει τη σελίδα, και να την κατατάξει στην κατάλληλη κατηγορία.

8που ασχολείται με τον καθορισμό ενός πρότυπου μοντέλου για την αποθήκευση γεωγραφικών πληροφοριών από πολλαπλές πηγές που επιτρέπουν η αποδοτική παραγωγή εξατομικευμένων γεωγραφικών λεξικών. Το GazDB χωρίζει τα ονόματα από τα χαρακτηριστικά αποθηκεύοντας τις σχέσεις μεταξύ τους. Τα γεωγραφικά ονόματα αποθηκεύονται με μια ποικιλία μορφών που να επιτρέπουν διεθνοποίησης ¹⁸ⁿ και την πολλαπλότητα της ονομασίας. Τα γεωγραφικά χαρακτηριστικά γνωρίσματα είναι ταξινομημένα κατά μήκος διαφόρων αξόνων που διευκολύνουν την επιλογή και το φιλτράρισμα. Ο σκοπός του MetaCarta GazDB είναι να παράσχει τόσο χώρο όσο και ενισχυτικούς μηχανισμούς για αποθήκευση, διατήρηση, και να εξαγωγή όλων όσων ξέρουμε για την συλλογή γεωγραφικών οντοτήτων μας.

Τέλος, μια πολύ ενδιαφέρουσα και πρόσφατη υπηρεσία ανοιχτού κώδικα, είναι ο μετατροπέας RSS σε GeoRSS ⁸. Με χρήση επεξεργασίας φυσικής γλώσσας (για την

⁴ **διεθνοποίηση.** Είναι η διαδικασία προσαρμογής ενός προϊόντος ή τοποθεσίας του Ιστού έτσι ώστε να μπορεί να χρησιμοποιήσει διαφορετικές γλώσσες και πολιτισμικές συμβάσεις χωρίς να απαιτείται εκ νέου σχεδιασμός. Η διεθνοποίηση πραγματοποιείται στο επίπεδο του σχεδιασμού του προγράμματος και ένα από τα βασικά του μελήματα είναι ο κώδικας να τρέχει αλάνθαστα σε μια πλατφόρμα με άλλη γλώσσα.



Σχήμα 2.2 : Γεωγραφική αναζήτηση με την υπηρεσία της Metacarta

εύρεση της πληροφορίας) και τεχνικών μηχανικής μάθησης (για να βελτιώνεται διαρκώς, αλληλεπιδρώντας με τους χρήστες), και βρίσκει γεωγραφική πληροφορία μέσα σε RSS feeds. Με εφαρμογή κυρίως σε ειδησεογραφικές ιστοσελίδες, παρουσιάζει τελικά έναν παγκόσμιο χάρτη, που δείχνει τα κατά τόπους γεγονότα. Έτσι η αναζήτηση και ανάγνωση των ειδήσεων μπορεί να γίνει με βάση γεωγραφικά κριτήρια.

2.3.2 Ερευνητικές Προτάσεις

Έχοντας κάνει μια μικρή επισκόπηση του εμπορικού τοπίου, ας δούμε και μερικές πρόσφατες ερευνητικές προτάσεις.

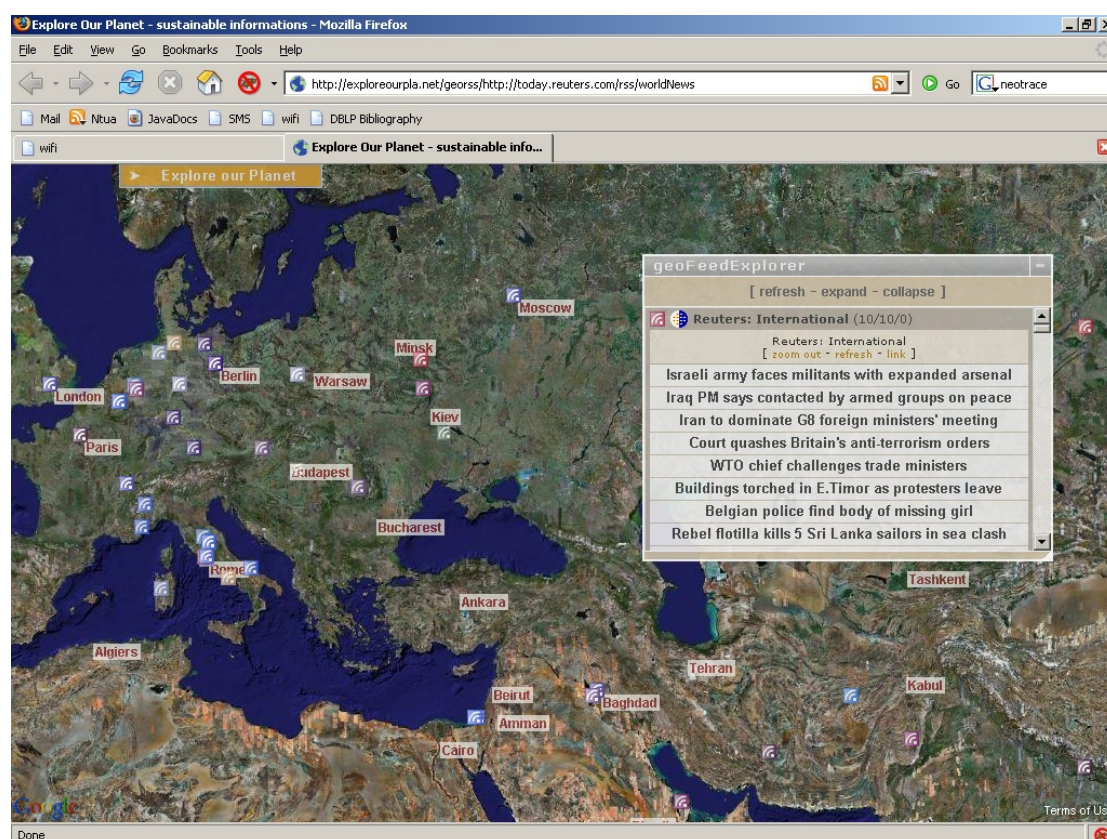
Στο 88 γίνεται εξαγωγή γεωγραφικής πληροφορίας από ιστοσελίδες, προκειμένου να ταξινομηθούν ως προς την γεωγραφική θέση τα αποτελέσματα μιας αναζήτησης. Δίνεται βαρύτητα και ανάλογα με την τοπικότητα μιας σελίδας (π.χ. μια σελίδα που αναφέρεται στα Θεωδωριανά Άρτησι¹⁰ αναμένεται να ενδιαφέρει λιγότερο κόσμο από μια που αναφέρεται στα Βαλκάνια γενικά).

Στο 89 παρουσιάζεται το Web-a-Where, ένα σύστημα που συνδέει μια ιστοσελίδα με την γεωγραφία. Το Web-a-Where εντοπίζει τις αναφορές σε τοποθεσίες και καθορίζει τη θέση στην οποία αναφέρεται κάθε μια. Επιπλέον, ορίζει σε κάθε σελίδα μία γεωγραφική περιοχή – την τοποθεσία δηλαδή στην οποία η σελίδα αναφέρεται συνολικά. Η διαδικασία της σηματοδότησης είναι απλή και γρήγορη, αφού στόχος είναι να εφαρμοστεί σε έναν μεγάλο αριθμό ιστοσελίδων και να διευκολύνει μια ποικιλία εφαρμογών που εξαρτώνται από την

εκάστοτε τοποθεσία. Τα πειράματα έδειξαν ότι στην περίπτωση των μεμονωμένων γεωγραφικών ονομάτων σε μια ιστοσελίδα η αντιστοίχιση τους με τις σωστές συντεταγμένες αγγίζουν το 80%, ενώ η αντιστοίχιση του γενικού περιεχομένου της ιστοσελίδας το 91%.

Στο κλασικό 8 περιγράφεται μια μηχανή γεωγραφικής πλοήγησης στο διαδίκτυο, για την οποία έχει προηγουμένως εξαχθεί γεωγραφική πληροφορία από αυτό. Προκύπτει έτσι ένας εναλλακτικός τρόπος πλοήγησης στο διαδίκτυο (βλ. Σχήμα 2.4).

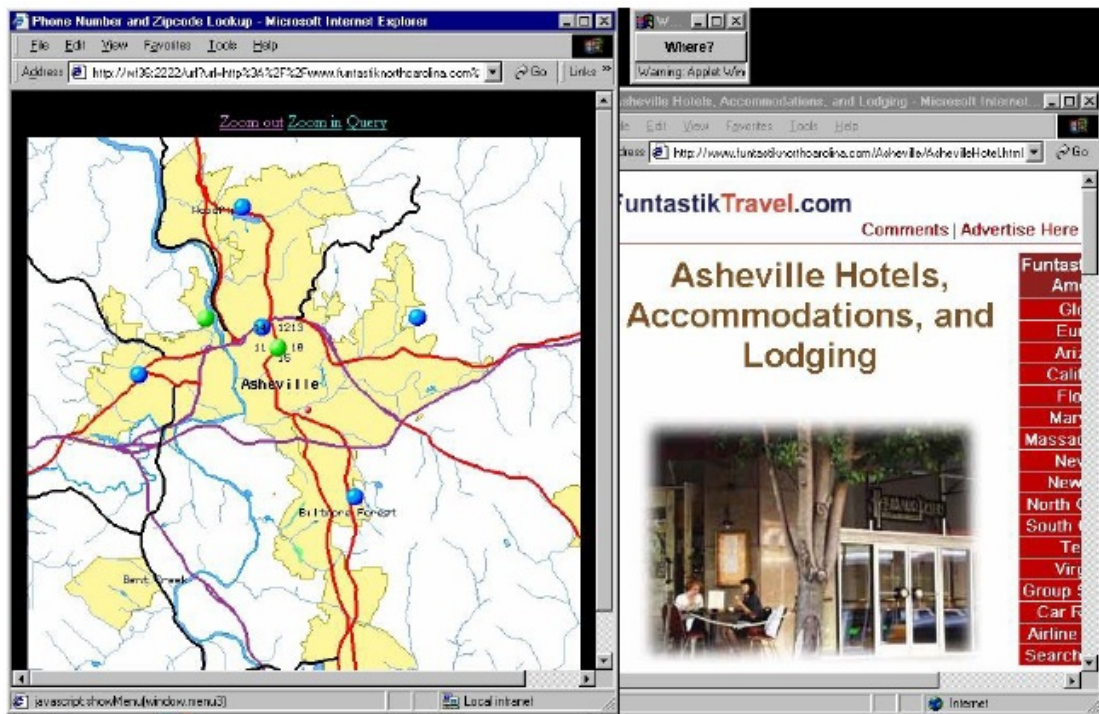
Στο 8προτείνεται η χρήση της γεωγραφικής πληροφορίας που διατίθεται στον παγκόσμιο ιστό για την πλήρωση με δεδομένα μιας χωρικής Β.Δ. (για παράδειγμα, στο περιβάλλον μιας υπηρεσίας βασισμένης στη θέση - Location Based Service). Εναλλακτικά, η πληροφορία μπορεί να χρησιμοποιηθεί για την αρχικοποίηση (bootstrapping) της βάσης.



Σχήμα 2.3: Διαβάζοντας τα νέα της μέρας με το RSS to GeoRSS converter 8

Τέλος, στο 8περιγράφεται μέθοδος υπολογισμού του γεωγραφικού εύρους μιας ιστοσελίδας, ενώ υλοποιείται και μια μηχανή γεωγραφικής αναζήτησης. Η μηχανή αυτή αξιοποιείται και στον τομέα της εξατομικευμένης αναζήτησης (personalised search), όπου η γεωγραφική θέση του χρήστη θεωρείται ως στοιχείο του προφίλ του, και άρα η εγγύτητα της γεωγραφικής θέσης μιας σελίδας σε αυτήν του χρήστη αποτελεί κριτήριο για την ιεράρχηση της σελίδας στην λίστα αποτελεσμάτων.

Σχήμα 2.4 : Γεωγραφική πλοήγηση στο διαδίκτυο. Από το 8



2.4 Σκοπός της διπλωματικής

Στις προηγούμενες παραγράφους περιγράφηκαν διάφορες προτάσεις που αφορούν τον τομέα της γεωκωδικοποίησης. Όλες οι προτάσεις αυτές βασίζονται στην εξαγωγή γεωγραφικών πληροφοριών από ιστοσελίδες με την βοήθεια κάποιου εργαλείου αυτόματης γεωκωδικοποίησης. Κάποιες από αυτές, όπως στο 8, προχωρούν ακόμα πιο μακριά προσθέτοντας στην εφαρμογή τους και μια μηχανή γεωγραφικής αναζήτησης ή όπως η metacart που δημιουργεί γεωγραφικά λεξικά ιστοσελίδων.

Τα αποτελέσματα, παρόλο τις εξαιρετικές προσπάθειες που έχουν γίνει σε αυτόν τον τομέα, εξακολουθούν να περιέχουν αρκετά λάθη, κυρίως λόγω προβλημάτων που είναι μέχρι στιγμής αδύνατον να αντιμετωπιστούν πλήρως (βλ. παρ. 2.2.1).

Όλα τα παραπάνω αποτέλεσαν ερέθισμα για την δημιουργία του εργαλείου που περιγράφεται σε αυτή την διπλωματική εργασία. Το εργαλείο αυτό σχεδιάστηκε να προσφέρει ημιαυτόματη γεωκωδικοποίηση σε ιστοσελίδες όταν ο χρήστης του το ζητήσει. Χρησιμοποιεί έναν αυτόματο γεωκωδικοποιητή ιστοσελίδων, εμπνευσμένο από τις εφαρμογές που περιγράφηκαν προηγουμένως, αλλά περιλαμβάνει ακόμα και ένα συμπληρωματικό εργαλείο που προσφέρει την δυνατότητα στον χρήστη να βελτιώσει τα αποτελέσματα που παρήχθησαν από το αυτόματο τρόπο.

Σε αντίθεση με τις εφαρμογές που προαναφέρθηκαν, το εργαλείο προσφέρει και την απεικόνιση των γεωγραφικών πληροφοριών, που έχουν βρεθεί, στον χρήστη με τρόπο

ευδιάκριτο αφού τις υπογραμμίζει και τις αντιστοιχίζει με τον γεωγραφικό χάρτη που να τις περιγράφει.

Τελικά δημιουργείται μια ενιαία βάση δεδομένων με στοιχεία που προέρχονται από το αυτόματο εργαλείο γεωκωδικοποίησης καθώς τις διορθωμένες εκδόσεις των προηγούμενων. Η βάση αυτή μπορεί να αποτελέσει βοήθημα για την δημιουργία ενός γεωγραφικού ευρετηρίου ιστοσελίδων (**gazeteer**) του ελληνικού Ιστού, αφού τα στοιχεία που περιέχει μπορούν να θεωρηθούν πιο συνεπή από αυτά που θα παρείχε μόνη της η αυτόματη γεωκωδικοποίηση.

Αυτή η προσέγγιση , όπως είναι φανερό είναι ιδιαίτερα χρήσιμη σε ιστοσελίδες που το περιεχόμενο τους αλλάζει βραχυπρόθεσμα (π.χ. Βικιπαιδεία), αφού οι βελτιώσεις που κάνει ο κάθε χρήστης θα μπορέσουν να διατηρηθούν για μεγαλύτερο χρονικό διάστημα.

3

Αυτόματη

Γεωκωδικοποίηση

Το παρόν κεφάλαιο αποτελεί μια περιγραφή της διπλωματική εργασίας του σπουδαστή 8η οποία χρησιμοποιήθηκε για το μέρος της παρούσας διπλωματικής εργασίας που ασχολείται με την αυτόματη γεωκωδικοποίηση των ιστοσελίδων. Ο κώδικας που αναπτύχθηκε χρησιμοποιήθηκε σχεδόν αυτούσιος με κάποιες μόνο μικρές αλλαγές, ώστε να μπορέσει να προσαρμοστεί την εφαρμογή μας. Στην συνέχεια ακολουθεί μια συνοπτική αναφορά τόσο για τον τρόπο ανάπτυξης του εργαλείου αυτόματης γεωκωδικοποίησης, όσο και στα προβλήματα στα οποία δεν μπόρεσαν να αντιμετωπιστούν. Σε περίπτωση που ο αναγνώστης θελήσει περισσότερες πληροφορίες μπορεί να ανατρέξει στην 8.

3.1 Σκοπός της διπλωματικής

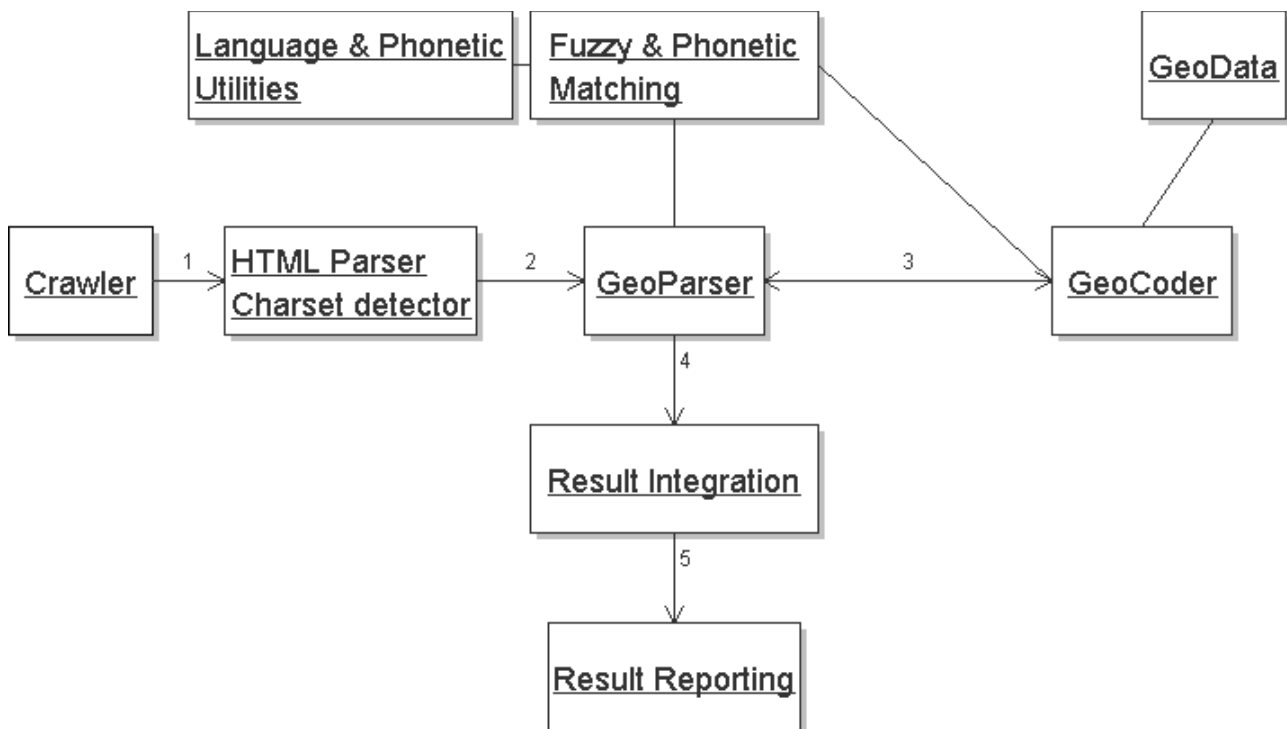
Όπως προαναφέρθηκε, η διπλωματική αυτή εργασία αποτελεί ένα εργαλείο για την αυτόματη γεωκωδικοποίηση ιστοσελίδων. Ο κύριος λόγος για τον οποίο δημιουργήθηκε είναι για την δημιουργία ενός γεωγραφικού ευρετηρίου του ελληνικού ιστοχώρου, δηλαδή για την απόδοση των γεωγραφικών συντεταγμένων σε ελληνικές ιστοσελίδες με βάση το τι περιγράφουν (το περιεχόμενό τους). Η επικέντρωση όμως στον ελληνικό κυβερνοχώρο (ο οποίος άλλωστε πάσχει από ένδεια τέτοιων υπηρεσιών) δημιουργεί αρκετά προβλήματα που θα εξεταστούν στην συνέχεια.

Αποτελεί μια λύση σε προβλήματα αναζήτησης ιστοσελίδων με βάση την γεωγραφική πληροφορία που περιέχουν και άλλων ανάλογων προβλημάτων. Είναι πλέον γνωστό, ότι πολλές σελίδες περιέχουν πλήθος γεωγραφικών πληροφοριών, ή έχουν έντονη γεωγραφική

χροιά. (Σύμφωνα με τον 8, σχεδόν 10% των σελίδων περιέχουν προφανή γεωγραφικά στοιχεία).

3.2 Αρχιτεκτονική Συστήματος

Ο σχεδιασμός της αρχιτεκτονικής του συστήματος, βασίστηκε πάνω στην κοινώς αποδεκτή μέθοδο της εξαγωγής γεωγραφικής πληροφορίας οδηγούμενης από το geoparsing. Δηλαδή, οι ιστοσελίδες συλλέγονται ξεχωριστά, και στην συνέχεια δίνονται στον geoparser για επεξεργασία. Αυτός φροντίζει για την εξαγωγή γεωγραφικής πληροφορίας από το κείμενο, την οποία και παραδίδει στον γεωκωδικοποιητή για γεωκωδικοποίηση. Τέλος, μια μονάδα του φροντίζει για την ολοκλήρωση της πληροφορίας, και την παρουσίαση των αποτελεσμάτων. Σε ξεχωριστές μονάδες έχουν τοποθετηθεί, επίσης, οι μέθοδοι που αφορούν την *προσεγγιστική αναζήτηση* και το *προσεγγιστικό ταίριασμα* (βλ. παρ 3.4.2 και 3.4.3). Το σύστημα έχει την μορφή μιας αλυσίδας επεξεργασίας. Στο Σχήμα 3.5 εικονίζεται ένα σκαρίφημα της αρχιτεκτονικής του, καθώς και η τυπική σειρά ακολουθιών για την επεξεργασία μιας ιστοσελίδας.



Σχήμα 3.5-Αρχιτεκτονική Αυτόματου Γεωκωδικοποιητή

3.2.1 Geoparsing

Στο τομέα του *geoparsing* χρησιμοποιήθηκαν διάφορες ιδέες από το περιβάλλον επεξεργασίας φυσικής γλώσσας GATE, γίνεται χρήση διαδοχικών κανονικών γραμματικών (cascaded regular grammar transducers), σε συνδυασμό με την *προσεγγιστική αναζήτηση* λέξεων σε ευρετήριο. Κάθε επίπεδο κανονικής γραμματικής χρησιμοποιεί τα αποτελέσματα του προηγούμενου, και εντοπίζει πληροφορίες διαφορετικής υφής. Για παράδειγμα, το πρώτο επίπεδο παίρνει ως είσοδο χαρακτήρες, και βγάζει λέξεις, αριθμούς, κενά, πιθανά τοπωνύμια κ.λ.π., ενώ το τελευταίο επίπεδο, δεχόμενο ως είσοδο ταχυδρομικούς κώδικες, τοπωνύμια, τηλέφωνα κ.ο.κ. βγάζει ως αποτελέσματα ολοκληρωμένες διευθύνσεις. Η *προσεγγιστική αναζήτηση* που αναφέρθηκε, χρησιμοποιείται για τον εντοπισμό πιθανών τοπωνυμίων στο κείμενο.

Παράλληλα, η πληροφορία που ανιχνεύεται, κανονικοποιείται και τεχνολογείται (*standardised & parsed*) (π.χ. η "οδ. Αγ.Σώστη 3, Χαλάνδρι 231-35" γίνεται {τύπος: "οδός", οδός: "Αγίου Σώστη", αρ: 3, περιοχή:"Χαλάνδρι", ΤΚ:"23135"}). Η κανονικοποίηση και ο διαχωρισμός της πληροφορίας σε πεδία, είναι απαραίτητα για την μετέπειτα γεωκωδικοποίησή της. Τέλος, όσον αφορά τα πιθανά τοπωνύμια που ευρίσκονται, γίνεται χρήση μιας ιδέας του Brill που περιγράφεται στο 8.

3.2.2 Γεωκωδικοποίηση - Geocoding

Η *γεωκωδικοποίηση* είναι η διαδικασία η οποία ακολουθείται ώστε μια γεωγραφική πληροφορία να μετατραπεί σε ένα ζεύγος ή τριάδα συντεταγμένων.

Για την πραγματοποίηση αυτής της διαδικασίας αρχικά συγκεντρώθηκαν επαρκή γεωγραφικά δεδομένα για τον ελλαδικό χώρο. Έπειτα, αναπτύχθηκαν αλγόριθμοι για την προσεγγιστική αναζήτηση όσων συμβολοσειρών χρειάζεται να αναζητηθούν (τοπωνύμια, διευθύνσεις κ.λ.π.). Ακολούθως, μια απλή λογική, με βάση την οποία γεωκωδικοποιείται η γεωγραφική πληροφορία, και της αντιστοιχίζεται ένας *συντελεστής βεβαιότητας*, και ένα ελάχιστο περιβάλλον ορθογώνιο. Ο *συντελεστής βεβαιότητας* είναι, αριθμητικά, η βεβαιότητα ότι το αντικείμενο που εξετάζεται πρόκειται περί γεωγραφικής πληροφορίας και κατά πόσο πως η γεωγραφική θέση αυτής είναι η ζητούμενη. Το δε ελάχιστο περιβάλλον ορθογώνιο περιγράφει την έκταση που έχει η θέση της πληροφορίας (δηλ. Το γεωγραφικό της εύρος).

3.2.2.1 Παράδειγμα γεωκωδικοποίησης

Για παράδειγμα, ας θεωρήσουμε μια διεύθυνση: Εφόσον συνοδεύεται από όνομα περιοχής, ή Τ.Κ., θα αναζητήσουμε τα στοιχεία αυτά στους αντίστοιχους καταλόγους. Εάν η διεύθυνση βρίσκεται εντός Αττικής, θα αναζητήσουμε και το όνομα του δρόμου στον αντίστοιχο κατάλογο. Τέλος, θα ελέγξουμε ότι οι γεωγραφικές θέσεις που λαμβάνουμε από τα επιμέρους

στοιχεία συμφωνούν μεταξύ τους (δηλ. έχουν μη κενή τομή). Σε αντίθετη περίπτωση, μειώνουμε τον συντελεστή βεβαιότητας κατά ποσοστό ανάλογο της εντοπιζόμενης ασυμφωνίας.

3.2.2.2 Ολοκλήρωση αποτελεσμάτων

Τα αποτελέσματα που προκύπτουν συναθροίζονται και συνεκτιμώνται, με βάση τον συσχετισμό της γεωγραφικής και κειμενικής τοπολογίας της σελίδας. Δηλαδή, εξετάζουμε παράλληλα την θέση της κάθε πληροφορίας στο χάρτη, όπως και μέσα στο κείμενο.

3.2.3 Γεωδεδομένα - GeoData

Προηγουμένως αναλύθηκε η διαδικασία της γεωκωδικοποίησης. Αναλύθηκε η προσέγγιση που χρησιμοποιήθηκε για την γεωκωδικοποίηση και καταδείχθηκε η ανάγκη ύπαρξης μεθόδων *προσεγγιστικής αναζήτησης*. Στο σημείο αυτό απλά είναι απαραίτητη η αναφορά στο τι είναι τα γεωδεδομένα, τι προβλήματα δημιουργήθηκαν και πως συλλέχτηκαν. Τα δεδομένα αυτά είναι λοιπόν, στην απλούστερη των περιπτώσεων, κατάλογοι αντιστοίχισης τοπωνυμίων σε γεωγραφικές συντεταγμένες. Λόγω της ελληνικής πραγματικότητας, δημιουργούνται διάφορα προβλήματα όπως η έλλειψη γεωδεδομένων, οι ιδιαιτερότητες που έχουν να κάνουν με γεωγραφικά συστήματα συντεταγμένων, καθώς και την ποιότητα των δεδομένων.

3.3 Παραδείγματα

Στο σημείο αυτό παρατίθεται ένα παράδειγμα πάνω σε μια ιστοσελίδα της Βικιπαιδείας για να δούμε τα αποτελέσματα που δίνει η αυτόματη γεωκωδικοποίηση. Για την απεικόνιση των αποτελεσμάτων ο συγγραφέας δημιουργεί ένα αρχείο τύπου HTML και ένα KML, όμως λόγω της δυσμορφίας που παρουσιάζει αυτό HTML αρχείο τα αποτελέσματα δεν είναι ιδιαίτερα εμφανής για την εφαρμογή μας. Έτσι θα χρησιμοποιήσουμε την απεικόνιση που προτιμήσαμε για την δική μας εφαρμογή, χωρίς βέβαια την επεξεργασία των αποτελεσμάτων, καθώς και το KML αρχείο.

3.3.1 Παράδειγμα «Κόρινθος»

Στο σχήμα Σχήμα 3.6 παρουσιάζονται τα αποτελέσματα από την αυτόματη γεωκωδικοποίηση. Είναι φανερό ότι τα αποτελέσματα δεν είναι απολύτως σωστά. Το αυτόματο έχει υπογραμμίσει λέξεις όπως κοινά, νομός ή έχει υπογραμμίσει μέρος μόνο των φράσεων που αποτελούν γεωγραφική πληροφορία π.χ. Κόρινθος αντί αρχαία Κόρινθος ή τέλος δεν τις έχει υπογραμμίσει καθόλου π.χ. Ακροκόρινθο. Οι λόγοι για τα παραπάνω

αναλύονται σε επόμενη ενότητα. Στο Σχήμα 3.6 φαίνονται με διαφορετικούς χρωματισμούς τα διάφορα λάθη που προέκυψαν.

Κόρινθος
 Από τη Βικιπαίδεια, την ελεύθερη εγκυκλοπαίδεια

Η **Κόρινθος** είναι πόλη και ο τρίτος σημαντικός **λιμένας** της Πελοποννήσου. Είναι η έδρα του ομώνυμου Δήμου, η **πρωτεύουσα του νομού Κορινθίας** και η μεγαλύτερη πόλη του νομού με 29.787 κατοίκους κατά την απογραφή του 2001 (36.565 ολόκληρος ο Δήμος).

Πίνακας περιεχομένων (Απόκρυφά)

- 1 Διοικητική διαίρεση του δήμου
- 2 Περιγραφή
- 3 Εξωτερικοί σύνδεσμοι
- 4 Φωτογραφικό υλικό

Διοικητική διαίρεση του δήμου [Επεξεργασία]

Ο δήμος περιλαμβάνει 5 δημοτικά διαμερίσματα:

- Κορίνθου (29.787 κάτοικοι)
- **Αρχαία Κορίνθος** (2.578 κάτοικοι)
- Εξαμίλιων (2.224 κάτοικοι)
- Ψαλοκέρρα (1.215 κάτοικοι)
- Σολιμού (751 κάτοικοι)

Περιγραφή [Επεξεργασία]

Η πόλη περιτοχίζεται από τις παραθαλάσσιες περιοχές του **Λεχαιού (σταθμό) Κεραιών** και τις περιοχές Εξαμίλιων και **Αρχαίας Κορίνθου**.

Η γεωγραφική της θέση είναι στην παραθαλάσσια περιοχή ανατολικά του **Κορινθιακού κόλπου** κοντά στην ζεύξη με τον Σαρωνικό.

Περιβάλλεται από το βουνό **Ονεία** και **Ασκαρύνθη**, όπου κατά την αρχαιότητα βρισκόταν η αρχαία πόλη της Κορίνθου.







Η Κόρινθος βρίσκεται στον οδικό άξονα **Αθήνα - Πάτρα**, με την πρόσβαση στην πρωτεύουσα να είναι εύκολη για τους κατοίκους της είτε μέσω της σύγχρονης εθνικής οδού είτε μέσω του **πρασσιτικού σιδηροδρόμου** που άρχισε να λειτουργεί το 2005.

Από το ακαδημαϊκό έτος 2003-2004 λειτουργεί στην Κόρινθο το **Τμήμα Κοινωνικής και Εκπαιδευτικής Πολιτικής** της Σχολής Κοινωνικών Επιστημών του Πανεπιστημίου Πελοποννήσου.

Εξωτερικοί σύνδεσμοι [Επεξεργασία]

- Ο ομότιμος οδηγός για την Κορίνθια
- <http://www.korinthos.gr>
- <http://www.e-korinthiakos.gr>
- <http://www.uop.gr/departments/depart8/index.shtml>
- Φωτογραφίες Κορίνθου
- Ζωντανή εικόνα (live webcam) από τον ναό του Απόλλωνα

Φωτογραφικό υλικό [Επεξεργασία]

			
Άποψη της Κορίνθου από το σιδηροδρομικό σταθμό σε επιστολικό δελτάριο των αρχών του 20ου αι.	Πύργος της Κορίνθου	Η γέφυρα του Ισθμού της Κορίνθου σε επιστολικό δελτάριο των αρχών του 20ου αι.	Πύλος του Απόλλωνα στην Αρχαία Κορίνθος
			
Ο τύραννος της Κορίνθου, Περικλής (περίπου 627-585 π.Χ.).	Η ναός του Αγίου Παύλου στην Κόρινθο σε επιστολικό δελτάριο των αρχών του 20ου αι.		

Σχήμα 3.6- Γεωγραφικό Εύρος της ιστοσελίδας

- Οι γεωγραφικές οντοτητές που δεν έχουν βρεθεί
- Όσες δεν έχουν βρεθεί πλήρως
- Λέξεις και φράσεις που έχουν βρεθεί λανθασμένα σαν γεωγραφικές οντότητες

Παρόλο που υπήρχαν αρκετά λάθη στα αποτελέσματα στο Σχήμα 3.7 παρατηρούμε έπειτα από συνεκτίμηση των αποτελεσμάτων, προέκυψε η θέση που φαίνεται με ορθογώνιο σχήμα στον χάρτη για την ιστοσελίδα. Βλέπουμε όμως ότι έχουν βρεθεί αποτελέσματα στην Μακεδονία, στον Βόλο, στα Δωδεκάνησα και στην Κρήτη, Παρόλο που το κείμενο δεν περιέχει τοποθεσίες σε αυτές της περιοχές. Οι τοποθεσίες αυτές είτε ανήκουν σε αυτές τις λέξεις που δεν αποτελούν γεωγραφική πληροφορία και βρέθηκαν από τον αυτόματο γεωκωδικοποιητή είτε σε κάποιες περιοχές έχουν δοθεί λάθος συντεταγμένες.



Σχήμα 3.7– Συνολική απεικόνιση τοποθεσιών

3.3.2 Παράδειγμα «Λεόντειο Αχαΐας»

Το παράδειγμα αυτό αποτελεί την γεωκωδικοποίηση της ιστοσελίδας της Βικιπαιδείας για το Λεόντειο Αχαΐας. Όπως είναι φανερό και από το Σχήμα 3.8 που απεικονίζει την ιστοσελίδα ύστερα από την αυτόματη γεωκωδικοποίηση, υπάρχουν αρκετές παραλήψεις όπως το “Λεόντειο”, Παναχαϊκός κ.τ.λ. Επίσης έχουν βρεθεί κάποιες λέξεις και φράσεις που δεν αποτελούν γεωγραφική πληροφορία, για παράδειγμα “πολλοί”, “Κατά”, “Απόστολος Ανδρέας”, κ.τ.λ. Και εδώ όπως και στο προηγούμενο παράδειγμα παρουσιάζουμε τα αποτελέσματα σημειωμένα με διαφορετικούς χρωματισμούς.

Βοήθεια · Τι είναι η Βικιπαίδεια · Βικιπαίδεια:Κανόνες συμπεριφοράς · Βικιπαίδεια:Αποστολή · Δεν είναι απαραίτητη αλλά προσαρμόζει [παραδείγματα](#)

Λεόντιο Αχαΐας

Από τη Βικιπαίδεια, την ελεύθερη εγκυκλοπαίδεια

Το Λεόντιο ονομάζεται ανεξάρτητη κοινότητα της ορεινής Αχαΐας ("Καποδιστριακή κοινότητα"). Μέχρι το 1914 ήταν δημοτικό διαμέρισμα του τότε Δήμου Φαρρών, και από τότε ανεξάρτητη κοινότητα. Βρίσκεται σε ύψος 750-800 μέτρων στους πρόποδες του **Παναχαϊκού** Πλησίον του χωριού υπάρχουν πολλές πηγές με κρυστάλλινα, πεντακάθαρα νερά καθώς και ένας παραπόταμος του Σελινούτα.

Πίνακας περιεχομένων [[Απεκκωμην](#)]

- 1 Ονομασία
- 2 Ιστορία
- 3 Λαϊκές παραδόσεις
- 4 Αναφορές

Ονομασία

Το σημερινό χωριό διατηρεί το όνομα της αρχαίας πόλης που υπήρχε εκεί.

Ιστορία

Διαβάστε για το αρχαίο **Λεόντιο** [Εδώ](#)

Ανασκαφές που έγιναν κατά τη δεκαετία του 1930 έφεραν στο φως τάφους και ευρήματα της Μυκηναϊκής εποχής 14ος-12ος αι. π.Χ. Υπάρχουν επίσης ίχνη ακρόπολης στο εκκλησάκι του Αγίου Κωνίνου που είναι ορατά ως σήμερα γύρω στον 3ο αι π.Χ. Λίγο πιο κάτω **επιπίσταναν** τα θεμέλια πρωτοχριστιανικής βασιλικής μεγάλων διαστάσεων πάνω στο πλάτωμα των οποίων είναι κτισμένος ο ναός του **Αγίου Ανδρέα** (4ος-5ος αι μ.Χ). Αυτά τα στοιχεία υποδηλώνουν συνεχή και έντονη ανθρώπινη παρουσία στην περιοχή, δυστυχώς όμως η έντονη διαβόηση λόγω των πολλών υδάτων, η ενκατάλειψη και οι περίοδοι αναστατώσεων του μεσαιώνα λήνα πόσινματα δίσωσαν από την αλλοτινή ακμή της περιοχής.

Το Λεόντιο είχε δημιουργήσει και αποικία ^[1] στην **Χίο**

Το Λεοντόποδο

Κατά τον Πλίνιο στο Λεόντιο υπήρχε μία πηγή ^[2] που οποιος έπινε νερό πέθανε μετά από τρεις ημέρες.

Στη σύγχρονη ιστορική σημαντικά γεγονόσ, που έλαβε χώρα στα περίχωρα του χωριού είναι η γνωστή **Μάχη του Λεοντίου** ανάμεσα σε ομάδα ανταρτών του ΕΛΑΣ και του 3ου συντάγματος "εθικού πολέμου" της παλικής Μεραρχίας Πεζομάχων υπό τον ταγματάρχη Γκασιόρα, τον επονομαζόμενο και "ανταρτοφάγο" Έπικεφαλής των ανταρτών ήταν ο καπετάνιος Μίχας. Στις 14 Ιουλίου του 1943 ξεκίνησε η μάχη με ενέδρα των ανταρτών. Στη μάχη που ακολούθησε οι έλληνες πατριότες κυριολεκτικά διέλυσαν τις ιταλικές κατοχικές δυνάμεις οι οποίες έχασαν γύρω στους εβδομήντα άνδρες όλο τον οπλισμό και τα υποζύγια τους. Με τα υλικά αυτά εισυθίθηκε αποφασιστικά η μαχητική ικανότητα του ΕΛΑΣ της βόρειας Πελοποννήσου. Αργότερα ως αντίποινα το χωριό κέρχκε από τις κατοχικές δυνάμεις που σκότωσαν τουλάχιστον δέκα χωρικούς. Πολλοί κάτοικοι του Λεοντίου συμμετείχαν ενεργά στον εφεδρικό ΕΛΑΣ και τις οργανώσεις του ΕΑΜ παρέχοντας ουσιαστική υποστήριξη στον απελευθερωτικό αγώνα τα χρόνια της κατοχής 1941-1944.

Λαϊκές παραδόσεις

Απο τα Λεόντια πέρασε ο **Απόστολος Ανδρέας**

^[3] αφού περιπλανήθηκε κυνηγημένος απο τους Ρωμαίους έφτασε στο χωριό, Αλλά οι κάτοικοί του τον αντιμετώπισαν άσχημα μην ξέροντας ποιός είναι. Μάλιστα στην θέση **Πάτρα** υπάρχουν δύο πλάκες που λένε ότι εκεί πάτησε ο Άγιος, στην μία αποτυπώθηκε το παπούτσι του και στην άλλη το δεξί του ποδι αφού είχα χάση το παπούτσι.

Μία άλλη παράδοση λέει ^[4] ότι **Απόστολος** ρχόμενος εκεί κυνηγημένος, απο τους στρατιώτες που τον έψαχναν για να τον σταυρώσουν, συνάντησε έναν χωριανό που που έπερνε το χωράφι του και του λέει "αν δεις στρατιώτες και σε ρωτήσουν, που ίσα κρύφτηκα, να μην το μαρτυρήσεις, πως ήρθα απο εδώ, γιατί θέλουν να με σταυρώσουν". Ο χωριανός δεν ήξερε ποί ήταν και απόρησε όταν είδε ότι μπροστά έπερνε και πίσω ψύτρωνε, και πρόδωσε τον άγιο μαρτυρώντας τον στους στρατιώτες. Αυτοί τον έδεσαν σπθάγκωνα και τον πήγαν στην **Πάτρα** και τον σταυρώσανε.

Αναφορές

- ↑ Αδαμόντιος Κοραΐς Ατακτα
- ↑ Πλίνιος Ν HIST.XXXI.19
- ↑ **Στέφανος Ουμπόπουλος** *Ιστορία της Πόλεως των Πατρών απο αρχαιοτάτων χρόνων μέχρι το 1821, Τόμος Α', Αχαϊκές Εκδόσεις, Πάτρα 1998* *ISBN 960-7960-10-6*
- ↑ **Στέφανος Ουμπόπουλος** *Ιστορία της Πόλεως των Πατρών απο αρχαιοτάτων χρόνων μέχρι το 1821, Τόμος Α', Αχαϊκές Εκδόσεις, Πάτρα 1998* *ISBN 960-7960-10-6*

Δήμοι και κοινότητες του νομού [[αποδείξεις](#)]

- Δήμος Πατρέων
- Δήμος Αχαΐας
- Δήμος Ανδρούσσης
- Δήμος Ακρόπολης
- Δήμος Αρσανίου
- Δήμος Βρανισσίων
- Δήμος Βαλιδόχου
- Δήμος Βαλιδόχου
- Δήμος Ερνεού
- Δήμος Καλαβρύτων
- Δήμος Λαρισσού
- Δήμος Μεσσήνιων
- Δήμος Μόδας
- Δήμος Παιών
- Δήμος Παροίων
- Δήμος Ρίου
- Δήμος Σικυωνίων
- Δήμος Τριπόλεως
- Δήμος Πάτρων
- Κοινότητα Καλεντινίου
- Κοινότητα Λεοντίου

Κατηγορίες: [Κοινότητα Λεοντίου](#) | [Χωριά της Αχαΐας](#)

Σχήμα 3.8 - Αυτόματη Γεωκωδικοποίηση της ιστοσελίδας της Βικιπαίδειας για το Λεόντιο Αχαΐας

Στην συνέχεια ακολουθεί ένας γεωγραφικός χάρτης (Σχήμα 3.9) της Ελλάδος με σημειωμένες όλες τις τοποθεσίες που βρέθηκαν στην ιστοσελίδα. Παρατηρούμε, ενώ έχουν

βρεθεί αρκετά σημεία στην περιοχή της Αχαΐας, έχουν σημειωθεί, όπως και στο προηγούμενο παράδειγμα, και αρκετά σημεία που δεν αναφέρονται στο κείμενο της ιστοσελίδας.



Σχήμα 3.9 - Απεικόνιση Τοποθεσιών της Ιστοσελίδας Λεόντειο Αχαΐας

3.4 Προβλήματα

Στο σημείο αυτό κρίνεται σκόπιμο να αναφερθούμε στους λόγους στους οποίους τα αποτελέσματα που δίνει το εργαλείο της αυτόματης γεωκωδικοποίησης στα παραπάνω παραδείγματα δεν ήταν απολύτως σωστά.

3.4.1 Το πρόβλημα της ταυτότητας του αντικείμενου

Γενικά, μιλάμε για το πρόβλημα ταυτότητας αντικείμενου όταν υπάρχει δυσκολία να "καταλάβει" ένα σύστημα πως τουλάχιστον δύο πλειάδες σε μία ή περισσότερες σχέσεις αναπαριστούν την ίδια οντότητα.

Στην συγκεκριμένη εργασία, τα προβλήματα που δημιουργούνται λόγω του προηγούμενου, είναι η χρονοβόρα διαδικασία για τον καθορισμό των γεωδεδομένων και ότι αποτελούν τη

βάση στο συναφές πρόβλημα του προσεγγιστικού ταιριάσματος και της προσεγγιστικής αναζήτησης.

Για την λύση του έχουν προταθεί αρκετοί αποδοτικοί αλγόριθμοι, κανένας όμως δεν μπορεί να προσφέρει απόλυτος σωστά αποτελέσματα.

3.4.1.1 Αιτίες του προβλήματος ταυτότητας αντικειμένου

Ας δούμε, συνοπτικά, μερικές από τις αιτίες του προβλήματος ταυτότητας αντικειμένου:

- Λάθος τιμές πεδίων, ή πεδία που λείπουν, εξαιτίας σφαλμάτων κατά την εισαγωγή δεδομένων. Π.χ.

Γεράσιμος	Σπανοδημήτρης	2102252252	Αθήνα	Άρρεν
Άσιμος	Σπανοδημήτρης	2102222252	null	Άρρεν

- Διαφορά στην ονοματολογία (π.χ. σε μία βάση έχουμε την κωδικοποίηση true= =Φίλαθλος false= =Μη φίλαθλος, ενώ στην άλλη έχουμε Φ= =Φίλαθλος και ΜΦ= =Μη φίλαθλος). Αυτό το πρόβλημα μπορεί να λυθεί εύκολα με μεθόδους μετάφρασης σχημάτων, και δεν χρειάζεται να μας απασχολήσει ιδιαίτερα εδώ.

- Μη πλήρης πληροφορία, γιατί τα απαραίτητα δεδομένα δεν καταχωρήθηκαν ή δεν είναι διαθέσιμα (π.χ. μια δημοσκόπηση καταγράφει μόνο τα αρχικά των ονομάτων των συμμετεχόντων)

- Αλλαγή σε κάποιο πεδίο ενδεχομένως δεν μεταβάλλει την ταυτότητα του αντικειμένου (π.χ. ο Γεράσιμος του παραδείγματος μετακομίζει στην Λάρισα)

- (Υστερόβουλη) Ψευδής δήλωση στοιχείων: Με τις αποθήκες δεδομένων να χρησιμοποιούνται για ανίχνευση απατών, ξεπλύματος χρήματος, αλλά και παρακολούθησης πολιτών, είναι αναμενόμενο κάποιος να εισαγάγει ψευδή στοιχεία, ή επιτηδευμένα ανορθόγραφα στοιχεία.

3.4.2 Προσεγγιστικό ταιρίασμα συμβολοσειρών

Αναφέρεται, ουσιαστικά, στο αν δύο δοθείσες συμβολοσειρές είναι όμοιες. Εδώ, όπως και προηγουμένως έχουν προταθεί και χρησιμοποιηθεί αρκετοί αλγόριθμοι για την εξυγίανση των αποτελεσμάτων, όπως η κανονικοποίηση και τεχνολόγηση των λέξεων και η μετατροπή τους σε φωνητικό αλφάβητο, αλλά παρόλα αυτά δεν είναι αρκετά.

3.4.2.1 Το πρόβλημα των *Greeklish*

Αυτή η μορφή επικοινωνίας, γνωστή στους περισσότερους χρήστες του ελληνικού κυβερνοχώρου, συνίσταται στην γραφή ελληνικών λέξεων, μεταγραμμένων σε λατινικούς χαρακτήρες με αυθαίρετο τρόπο.

Η λύση στο πρόβλημα με τα greeklish δίνεται από το Ινστιτούτο Επεξεργασίας Λόγου, με μια εφαρμογή τόσο κομψή όσο και με εξαιρετικής ποιότητας αποτελέσματα, την All Greek To Me!.

3.4.3 Προσεγγιστική αναζήτηση συμβολοσειρών

Τελικός στόχος όλης της προηγούμενης δουλειάς είναι η προσεγγιστική αναζήτηση των συμβολοσειρών από έναν κατάλογο για την ανάκτηση κάποιων στοιχείων της που την χαρακτηρίζουν μοναδικά (π.χ. Id). Η απευθείας σύγκριση της αναζητούμενης συμβολοσειράς με όλες τις υπόλοιπες είναι υπολογιστικά ανέφικτη, και γι' αυτό απαιτείται κάποιος γρήγορος τρόπος "φιλτραρίσματος" των υποψηφίων ταιριασμάτων έτσι ώστε αφ' ενός να είναι απίθανο το να απορριφθεί εσφαλμένα κάποια συμβολοσειρά, αφ' ετέρου να απορριφθούν (ορθά) όσο το δυνατόν περισσότερες συμβολοσειρές. Χρησιμοποιήθηκε η αναζήτηση με βάση ένας αλγόριθμος αναζήτησης σε τρεις παραλλαγές, ανάλογα με το μέγεθος του καταλόγου αναζήτησης, του τύπου των δεδομένων που αναζητούνται, και της ταχύτητας που επιδιώκεται, το *ευρετήριο κατακερματισμού μεμονωμένων λέξεων*, το *ευρετήριο κατακερματισμού πολλαπλών λέξεων* και το *φωνητικό ευρετήριο*. Για περισσότερες πληροφορίες πάνω στους αλγόριθμους αναζήτησης ο χρήστης καλείται να ανατρέξει στις παραγράφους 4.4.5 – 4.4.7 της 8.

3.4.4 Το Πρόβλημα των Γεωδεδομένων

Στην παράγραφο 3.2.3 έγινε αναφορά για τα προβλήματα που μπορούν να προκληθούν από τα γεωδεδομένα. Ακολουθεί μια αναλυτική αναφορά σε αυτά, ώστε να γίνει σαφής η σημαντικότητα τους.

3.4.4.1 Έλλειψη δεδομένων

Καταρχήν, επισημαίνουμε την έλλειψη γεωγραφικών δεδομένων. Ενώ σε πολλές χώρες του εξωτερικού, διατίθενται δωρεάν και δημόσια γεωγραφικά δεδομένα υψηλής ποιότητας (λ.χ. Στις Η.Π.Α. διατίθεται, μεταξύ άλλων²², γεωκωδικοποιημένος κατάλογος οδών, πόλεων, ταχυδρομικών κωδίκων, κ.ο.κ.²³, στις περισσότερες χώρες διατίθεται τηλεφωνικός κατάλογος σε ηλεκτρονική μορφή, κ.λ.π.), στην Ελλάδα δεν υπάρχει κάποια κεντρική υπηρεσία διάθεσής τους. Ορισμένες υπηρεσίες (π.χ. ΕΛΤΑ, ΟΤΕ) προσφέρουν τα δεδομένα τους, αλλά μόνο για μεμονωμένα ερωτήματα στον εξυπηρετητή τους, κάνοντας την χρήση τους από προγράμματα δυσχερή. Συνεπώς, η ανάγκη για γεωδεδομένα καλύπτεται είτε καταφεύγοντας στην δαπανηρή λύση της αγοράς τους από εταιρείες του χώρου, είτε ψάχνοντας εναλλακτικούς τρόπους απόκτησής τους.

3.4.4.2 Ποιότητα δεδομένων

Αγνοώντας προς το παρόν την ένδειά τους, τα ελληνικά γεωδεδομένα υποφέρουν και από προβλήματα πληρότητας, ποιότητας και ακρίβειας. Σχετικά με την πληρότητα, επειδή ο πληθυσμός της χώρας μας είναι ανισομερώς κατανομημένος, είναι πολύ ευκολότερο ή/και φθηνότερο να βρεθούν γεωδεδομένα για μια πυκνοκατοικημένη περιοχή, π.χ. για τον νομό Αττικής (όπου, άλλωστε, το κόστος απόσβεσης της απόκτησης των σχετικών γεωδεδομένων καλύπτεται γρήγορα), απ' ό,τι για μια επαρχιακή. Η υποβέλτιστη ποιότητα των δεδομένων προκύπτει από την άντλησή τους από ετερόκλητες πηγές, πολλές από τις οποίες δεν διασφαλίζουν την ποιότητά των δεδομένων τους. Ακόμη, επειδή τα περισσότερα ελληνικά γεωδεδομένα διατίθενται από ξένες υπηρεσίες (π.χ. Την γεωγραφική υπηρεσία των ΗΠΑ), και άρα με λατινικούς χαρακτήρες, τίθεται ζήτημα μεταγραφής τους στα ελληνικά. Τέλος, η χρήση κατά το παρελθόν ποικίλων συστημάτων συντεταγμένων (Bessel, HATT κ.λ.π.), τα οποία τροποποιούνταν συνεχώς, αλλά και η έλλειψη εύχρηστων εφαρμογών μετατροπής συντεταγμένων μεταξύ του παγκόσμιου συστήματος αναφοράς WGS84, και του τοπικά χρησιμοποιούμενου ΕΓΣΑ87, έχει οδηγήσει αρκετά datasets γεωδεδομένων στο να περιέχουν σημαντικά σφάλματα ή μειωμένη ακρίβεια στις συντεταγμένες τους.

3.4.4.3 Καθαρισμός δεδομένων

Ο καθαρισμός δεδομένων είναι μια διαδικασία που αποσκοπεί στην βελτίωση της ποιότητας (και, άρα, της αξίας) ενός συνόλου δεδομένων (dataset). Αυτό επιτυγχάνεται με διάφορες τεχνικές, που συνήθως αποτελούν μείξη τυπικών μεθόδων και εξειδικευμένων, για την συγκεκριμένη περίπτωση, λύσεων. Ο κύριος στόχος της διαδικασίας είναι η εξασφάλιση της ορθότητας του κάθε δεδομένου (π.χ. μια τοποθεσία στην Ελλάδα δεν μπορεί να έχει συντεταγμένες που αναφέρονται στη Σιβηρία), της συνέπειας των δεδομένων ενός συνόλου (π.χ. μπορεί να μην είναι επιθυμητό δύο εγγραφές να αναφέρονται στην ίδια τοποθεσία), καθώς και της συνέπειας μεταξύ των διαφόρων συνόλων δεδομένων (π.χ. σε ένα σύνολο δεδομένων ο T.K. 35323 αντιστοιχίζεται σε μια περιοχή στη Μακεδονία, ενώ σε ένα άλλο σύνολο, μια οδός στη Λαμία έχει τον ίδιο T.K.). Ακόμη, η διαδικασία πρέπει να είναι οργανωμένη σε μορφή ροής, στην είσοδο της οποίας δίνονται τα ακατέργαστα δεδομένα, και από έξοδο της οποίας λαμβάνονται καθαρά. Ο λόγος για αυτή την αρχιτεκτονική είναι, απ' ενός η αυτοματοποίηση της διαδικασίας, απ' ετέρου η δυνατότητα αυξητικής (incremental) εισαγωγής νέων δεδομένων στο σύστημα. Ο ενδιαφερόμενος αναγνώστης παραπέμπεται στο 8για περισσότερα σχετικά με το γνωστικό αντικείμενο αυτό. Στην μεθεπόμενη παράγραφο, θα δούμε μια τυπική εφαρμογή καθαρισμού δεδομένων, στα πλαίσια ενός συνόλου γεωδεδομένων.

3.5 Η δική μας μετατροπή

Η προηγούμενη παράγραφος κάνει περισσότερο από σαφές ότι παρόλο την χρησιμοποίηση ευφών και ιδιαίτερα αποτελεσματικών αλγορίθμων τα λάθη είναι αναπόφευκτα. Έτσι, δημιουργήθηκε η ανάγκη για την δημιουργία ενός εργαλείου, το οποίο θα έχει την δυνατότητα να διορθώνει τα λάθη από την αυτόματη γεωκωδικοποίηση μιας ιστοσελίδας.

Τα λάθη αυτά μπορεί να είναι :

- Λανθασμένες συντεταγμένες σε μια γεωγραφική οντότητα
- Λανθασμένη επισήμανση λέξης ή φράσης ως γεωγραφική οντότητα
- Αναγνώριση τμήματος φράσης ως γεωγραφική οντότητα(Κόρινθος αντί για αρχαία Κόρινθος)
- Μη αναγνώριση λέξης ή φράσης ως γεωγραφική οντότητα

Το εργαλείο της αυτόματης γεωκωδικοποίησης που περιγράφηκε προηγουμένως δέχεται σαν είσοδο ένα σύνολο από ιστοσελίδες, τις επεξεργάζεται βρίσκοντας από το κείμενο που περιέχουν τις γεωγραφικές οντότητες και στη συνέχεια αντιστοιχεί γεωγραφικές συντεταγμένες σε κάθε ιστοσελίδα.

Ορμώμενοι από όλα τα παραπάνω δημιουργήσαμε ένα εργαλείο που θα βελτιώνει τα αποτελέσματα της γεωκωδικοποίησης σε ιστοσελίδες που το περιεχόμενο τους μένει σταθερό βραχυπρόθεσμα.

4

Ανάλυση Απαιτήσεων

Συστήματος

Στο κεφάλαιο αυτό ακολουθεί μια περιγραφή της λογικής πάνω στην οποία χτίστηκε το παρόν σύστημα. Γίνεται, επίσης, ανάλυση των διαφόρων κομματιών που το αποτελούν καθώς και του τρόπου που συνδέονται μεταξύ τους. Περισσότερες λεπτομέρειες για τις τεχνολογίες που χρησιμοποιήθηκαν υπάρχουν στο επόμενο κεφάλαιο.

4.1 Αρχιτεκτονική

Όπως ειπώθηκε και προηγουμένως ο σκοπός της παρούσας εργασίας είναι η κατασκευή ενός εργαλείου για την ημιαυτόματη γεωκωδικοποίηση των ιστοσελίδων που ο χρήστης έχει ανοίξει στον φυλλομετρητή του, καθώς επίσης, και για την απεικόνιση αυτών των γεωγραφικών πληροφοριών με τρόπο ευδιάκριτο προς τον χρήστη (υπογράμμιση, γεωγραφικοί χάρτες). Τέλος, το εργαλείο αυτό προσφέρει στους χρήστες την δυνατότητα διόρθωσης με διάφορους τρόπους των τυχόν λαθών που μπορεί να υπάρξουν από την διαδικασία της γεωκωδικοποίησης.

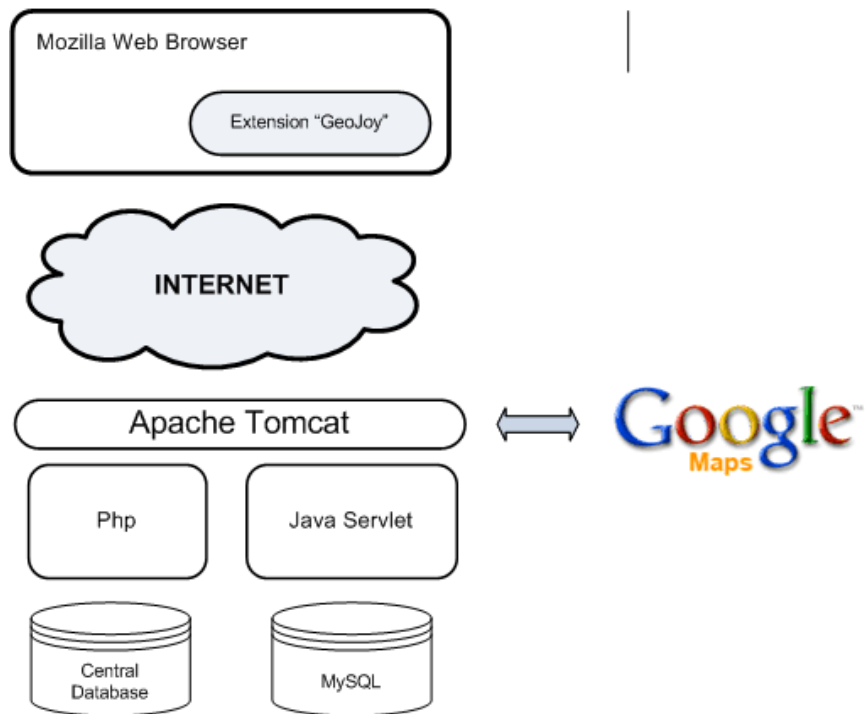
Είναι φανερό ότι η όλη η διαδικασία ξεκινάει όταν ο χρήστης ανοίξει μια ιστοσελίδα στον φυλλομετρητή και θελήσει να δει τις γεωγραφικές πληροφορίες που αυτό περιέχει. Επιθυμούμε το εργαλείο που θα δημιουργήσουμε να είναι ενσωματωμένο στον φυλλομετρητή, ώστε να είναι φιλικό προς τον χρήστη και αποδοτικό. Η τεχνολογία που έχει την δυνατότητα να προσφέρει όλα τα προηγούμενα είναι αυτή των λεγόμενων *extensions* για φυλλομετρητές. Τα *extension* είναι γραμμένα στην γλώσσα προγραμματισμού *JavaScript* που περιέχει συναρτήσεις για την τροποποίηση των ιστοσελίδων, κάτι που βοηθάει στην απεικόνιση των γεωγραφικών πληροφοριών. Επίσης, περιέχει εργαλεία για την επικοινωνία

με HTTP πρωτόκολλα, κάτι που βοηθάει στην διαδικασία τόσο της γεωκωδικοποίησης όσο και της διόρθωσης των λαθών, όπως θα δούμε και στην συνέχεια. Η ροή εκτέλεσης του αλγορίθμου βασίζεται, ουσιαστικά, στον ίδιο το extension, αφού αυτό είναι υπεύθυνο για την διεξαγωγή και το συντονισμό όλων των ενεργειών.

Η διαδικασία της γεωκωδικοποίησης αποτελείται από δύο μέρη, την γεωκωδικοποίηση εκτελείται από τον αυτόματο γεωκωδικοποιητή που περιγράψαμε στο κεφ. 3, και αυτή που εκτελείται χειρονακτικά, όπου ο χρήστης βελτιώνει τα αποτελέσματα που έχουν παραχθεί με την αυτόματη διαδικασία. Και στις δύο περιπτώσεις τα δεδομένα είναι αποθηκευμένα σε μια κεντρική βάση δεδομένων που βρίσκεται σε έναν εξυπηρετητή. Ο αυτόματος γεωκωδικοποιητής αποτελεί ένα αυτόνομο εργαλείο γραμμένο στην γλώσσα προγραμματισμού Java (παρ. 5.8) κάτι που κάνει την ενσωμάτωση του στο extension προβληματική στην απλότητα και στην αποδοτικότητα του εργαλείου μας. Έτσι προτιμήθηκε να χρησιμοποιηθεί η τεχνολογία των Java Servlets (βλ. παρ 5.7) όπου το αυτόματο απλά φιλοξενείται σε έναν εξυπηρετητή (στην περίπτωση μας τον Apache Tomcat) και εκτελείται μόνο όταν χρειάζεται.

Στην διαδικασία της απεικόνισης χρησιμοποιούνται κάποιοι χάρτες της εταιρείας Google που περιγράφουν αυτές τις γεωγραφικές πληροφορίες, με την βοήθεια της τεχνολογίας που προσφέρει το Google Map API (βλ. παρ. 5.3).

Το Σχήμα 4.10 αποτελεί το βασικό σκαρίφημα της αρχιτεκτονικής του συστήματος μας και απεικονίζει τον τρόπο που τα κομμάτια του συστήματος συνδέονται μεταξύ τους.



Σχήμα 4.10- Αρχιτεκτονική Συστήματος

4.2 Περιγραφή Λειτουργιών

Το εύρος των λειτουργιών χωρίζεται σε δύο βασικές λειτουργίες :

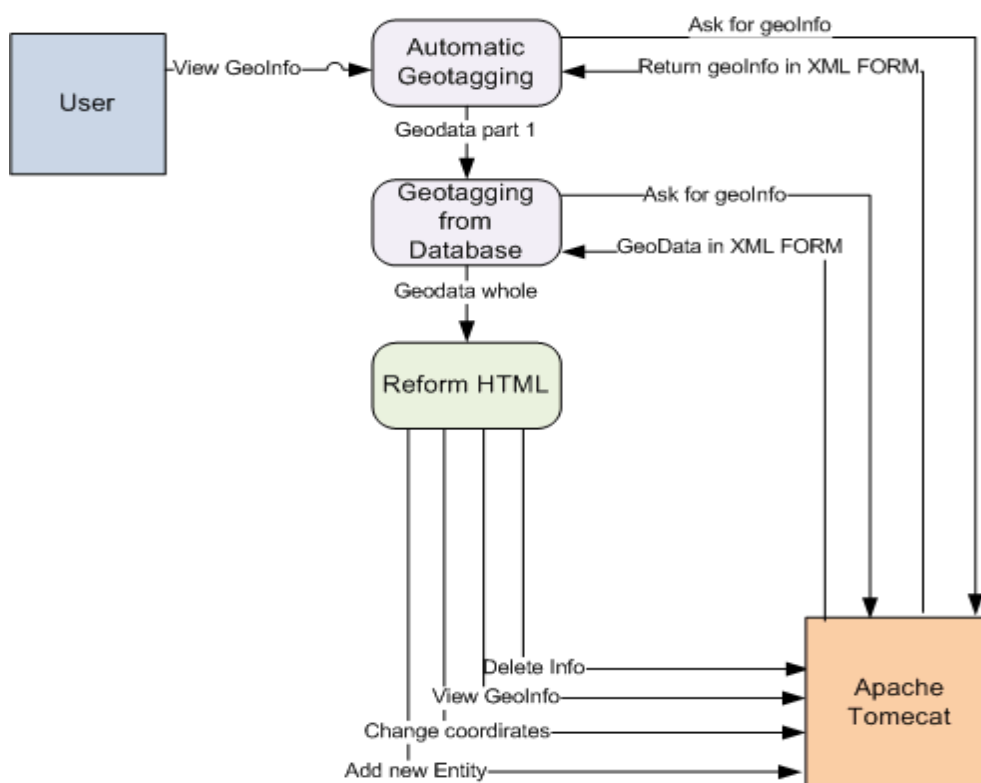
- την ανάκτηση της γεωγραφικής πληροφορίας από μια ιστοσελίδα καθώς και την παρουσίαση της στον χρήστη και
- τη διόρθωση λαθών ή την προσθήκη δεδομένων που ανακτήθηκαν λάθος ή παραλήφθηκαν λόγω της μη αποτελεσματικότητας του αλγορίθμου και την ανανέωση ή προσθήκη στην κεντρική βάση των τελευταίων για μελλοντική χρήση.

Όπως είναι φανερό από το Σχήμα 4.10 το σύστημα χωρίζεται σε δύο υποσυστήματα, το Browser-Extension, το Server όλα με προκαθορισμένες λειτουργίες.

4.2.1 Browser Extension (client side)

Αποτελεί, όπως ειπώθηκε και προηγουμένως, το βασικό υποσύστημα, αφού είναι εκείνο που περιέχει το κύριο menu της εφαρμογής δηλαδή αυτό ελέγχει πότε και με ποιόν τρόπο θα εκτελεστούν όλες οι λειτουργίες. Η κύρια εφαρμογή του συστήματος είναι η εύρεση και η απεικόνιση των γεωγραφικών πληροφοριών μιας ιστοσελίδας στον χρήστη. Το extension λαμβάνει τις γεωγραφικές πληροφορίες από τον αλγόριθμο αυτόματης εύρεσης γεωγραφικών πληροφοριών και από την ανάγνωση της κεντρικής βάσεως δεδομένων, τα συνθέτει και τα παρουσιάζει στον φυλλομετρητή τροποποιώντας το αντίστοιχο αρχείο HTML για να είναι ορατές στον χρήστη. Στην συνέχεια ο χρήστης αν εντοπίσει κάποιο λάθος στα προηγούμενα

αποτελέσματα έχει το δικαίωμα να τα τροποποιήσει(αλλάζοντας τις συντεταγμένες του αντίστοιχου χάρτη) ή ακόμη και να αφαιρέσει κάποια οντότητα που δεν αντιστοιχεί σε γεωγραφική πληροφορία με την βοήθεια ειδικών menu, δίνοντας έτσι εντολή στο υποσύστημα Server να κάνει τις απαραίτητες ενέργειες, ώστε την επόμενη φορά να απεικονιστούν τα νέα δεδομένα. Τέλος, αν κάποια πληροφορία δεν έχει βρεθεί, και παρατηρηθεί από τον χρήστη, ο τελευταίος δύναται να προσθέσει αυτή την πληροφορία. Όπως και στην προηγούμενη περίπτωση τα αποτελέσματα καταχωρούνται στην κεντρική βάση δεδομένων (database2) με τη βοήθεια του υποσυστήματος Server για την σωστότερη ανάκτηση της πληροφορίας της ιστοσελίδας μελλοντικά.

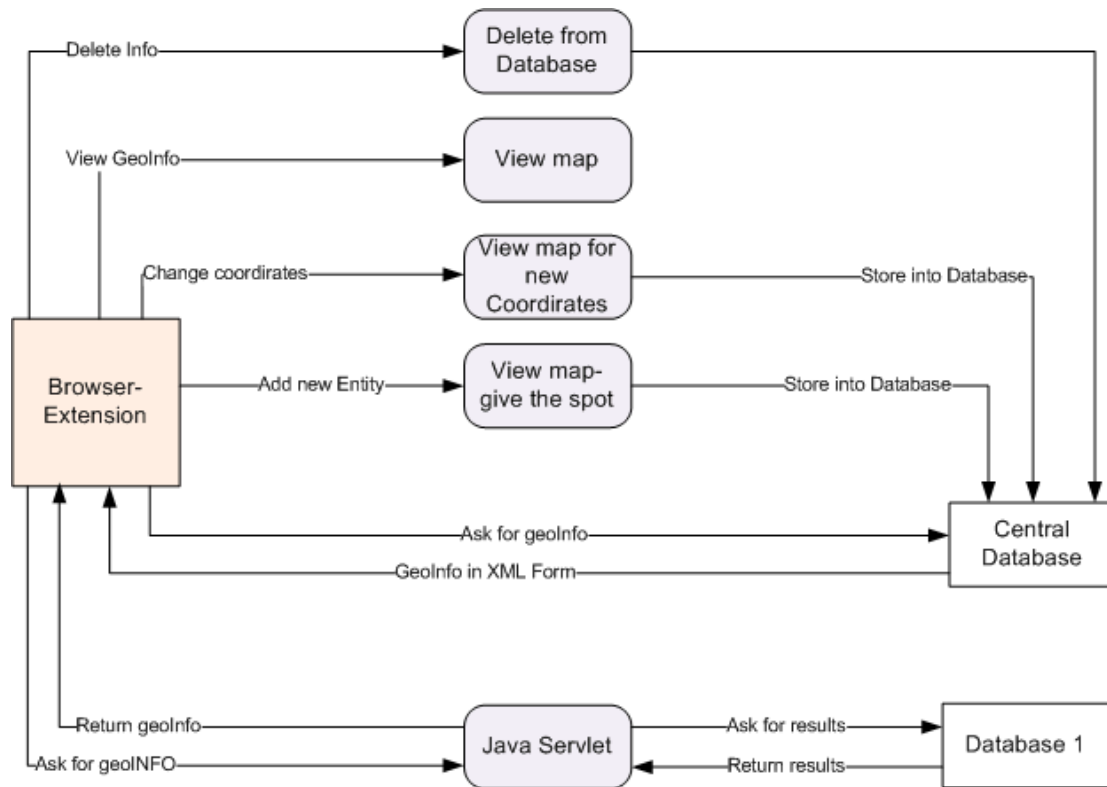


Σχήμα 4.11 – Υποσύστημα Browser-Extension

4.2.2 Server (server side)

Το παρόν υποσύστημα μπορεί να χωριστεί περαιτέρω σε δύο μέρη λόγω των εργασιών που εκτελούν αυτά.

Το πρώτο Server – Java Servlet κάνει χρήση αλγορίθμου που αναπτύχθηκε στην διπλωματική “Εξαγωγή Γεωγραφική Πληροφορίας από Ημιδομημένο Κείμενο” για την εύρεση των γεωγραφικών οντοτήτων της συγκεκριμένης ιστοσελίδας με χρήση της τεχνολογίας των Java Servlets. Είναι, επίσης, υπεύθυνο για την σωστή παράδοση των αποτελεσμάτων στο Browser Extension, τα οποία θα επεξεργαστούν και θα εμφανιστούν στην ιστοσελίδα.



Σχήμα 4.12 – Υποσύστημα Tomcat - Βάσεις Δεδομένων

Το Server-CentralDatabase περιέχει την κεντρική βάση δεδομένων του συστήματος, αφού σε αυτή αποθηκεύονται όλα τα τελικά δεδομένα(είτε από τον αλγόριθμο αυτόματης εύρεσης γεωγραφικής πληροφορίας είτε από τις βελτιώσεις του χρήστη, όπως εξηγήθηκε στην προηγούμενη παράγραφο) που χρησιμοποιούνται για την ανάκτηση της γεωγραφικής πληροφορίας σε μια ιστοσελίδα. Τέλος, ο Server είναι υπεύθυνος για την ανάκτηση των Google maps και της προσθήκης τους με τις σωστές παραμέτρους στην ιστοσελίδα καθώς και για την επικοινωνία του extension με την κεντρική βάση δεδομένων για την επεξεργασία των περιεχομένων της.

4.3 Ανάλυση Επιμέρους Τμημάτων

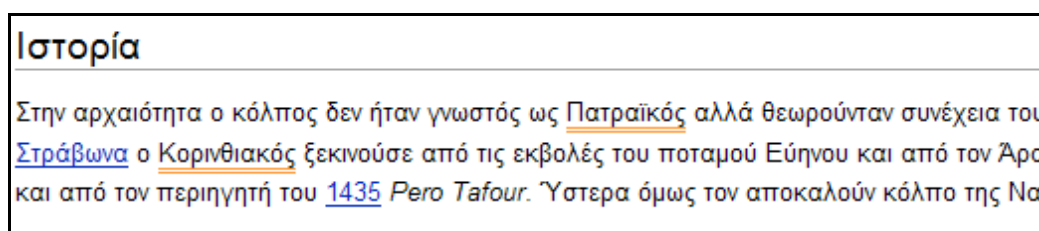
Από τα προηγούμενα έχει γίνει σαφές ότι η εφαρμογή χωρίζεται σε διάφορα επιμέρους κομμάτια που χρησιμοποιούν διαφορετικά είδη τεχνολογίας Ιστού, με κάθε ένα από αυτά να έχει έναν ξεχωριστό και ιδιαίτερα σημαντικό ρόλο για την λειτουργία του τελικού συστήματος. Στην συνέχεια ακολουθεί η περιγραφή της υλοποίησης αυτών των κομματιών καθώς και του τρόπου σύνδεσής τους.

4.3.1 User Interface - Browser Extension

Το κομμάτι αυτό του κώδικα είναι αυτό που είναι ουσιαστικά υπεύθυνο για τη ροή εκτέλεσης της εφαρμογής αφού είναι αυτό που περιέχει το user interface, δηλαδή είναι το μόνο κομμάτι με το οποίο ο χρήστης έρχεται σε άμεση επαφή. Η υλοποίηση του έχει γίνει με την βοήθεια της γλώσσας προγραμματισμού JavaScript.

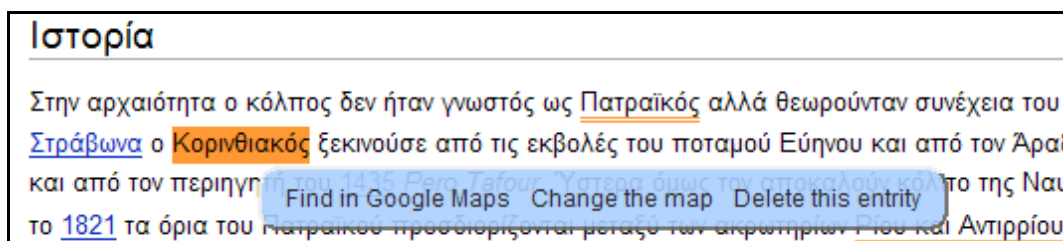
Η κύρια ενέργεια που εκτελείται είναι η απεικόνιση της γεωγραφικής πληροφορίας της εκάστοτε ιστοσελίδας στο χρήστη. Οι λέξεις ή οι φράσεις που αντιστοιχούν σε γεωγραφική πληροφορία υπογραμμίζονται (Σχήμα 4.13) και όταν ο κέρσορας σταματήσει πάνω τους εμφανίζεται ένα hover-over menu με τρεις επιλογές (Σχήμα 4.14):

- α) να δει ο χρήστης τον χάρτη στον οποίο αντιστοιχεί η γεωγραφική πληροφορία
- β) να δει το χάρτη και σε περίπτωση που οι συντεταγμένες του έχουν απεικονιστεί λανθασμένα να δώσει τις σωστές
- γ) αν η υπογραμμισμένη λέξη ή φράση δεν αποτελεί γεωγραφική πληροφορία να την διαγράψει από την βάση.



Σχήμα 4.13 – Απεικόνιση Γεωγραφικής Πληροφορίας

Για την απεικόνιση της πληροφορίας χρησιμοποιούνται οι καταχωρίσεις που υπάρχουν σε μια κεντρική βάση δεδομένων για την συγκεκριμένη ιστοσελίδα. Οι καταχωρίσεις αυτές αποτελούνται από δεδομένα που προέρχονται από τα αποτελέσματα του αυτόματου γεωκωδικοποιητή ιστοσελίδων, τις διορθωμένες εκδόσεις τους από τους χρήστες καθώς επίσης και δεδομένα που έχουν προσθέσει από αυτούς.



Σχήμα 4.14 – Απεικόνιση menu επιλογών

Η αυτόματη γεωκωδικοποίηση εκτελείται όταν δεν υπάρχει καμία καταχώριση για αυτήν την ιστοσελίδα στη κεντρική βάση δεδομένων από τον αυτόματο αλγόριθμο ή όταν το περιεχόμενο της ιστοσελίδας έχει αλλάξει σε σχέση με την προηγούμενη εκτέλεση της. Στην

περίπτωση αλλαγής του περιεχομένου όσες καταχωρίσεις δεν βρεθούν συμβατές με την νέα μορφή της ιστοσελίδας απορρίπτονται. Τόσο η κεντρική βάση δεδομένων όσο και το εργαλείο της αυτόματης γεωκωδικοποίησης (με την μορφή του Java Servlet) βρίσκονται σε ένα server (Apache Tomcat). Για την επικοινωνία τους λοιπόν με το extension χρησιμοποιήθηκε το πρότυπο AJAX(Asynchronous JavaScript and XML) και πιο συγκεκριμένα η τεχνική XmlHttpRequest. Το αντικείμενο XMLHttpRequest δημιουργήθηκε για να επιτρέψει την πραγματοποίηση ασύγχρονων αιτημάτων HTTP από το extension στον server. Αυτά τα αιτήματα επιστρέφουν τα αποτελέσματα σε μορφή XML, μια μορφή που παρείχε έναν εύκολο τρόπο επεξεργασίας πληροφοριών.

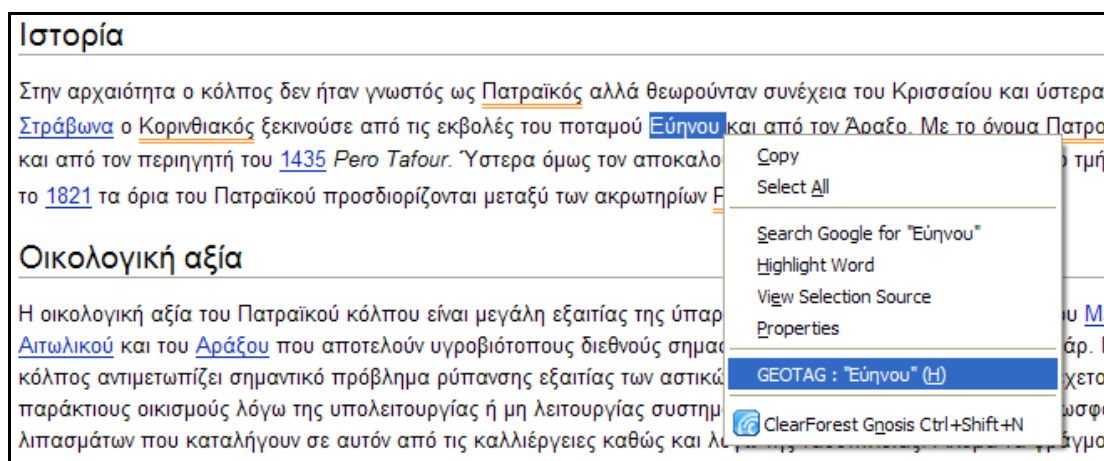
Στην περίπτωση του εργαλείου της αυτόματης γεωκωδικοποίησης, που είναι γραμμένο σε Java από τον κατασκευαστή, πραγματοποιήσαμε απλά μετατροπή του σε Java Servlet (για περισσότερες πληροφορίες 5.7), η επικοινωνία με το extension γίνεται απλά με την χρήση των συναρτήσεων του AJAX. Αντίθετα, στην περίπτωση επικοινωνίας με την κεντρική βάση δεδομένων χρειάζεται να υπάρξει η μεσολάβηση της PHP. Η PHP αποτελεί server-side scripting και έχει την δυνατότητα να χειρίζεται με αποτελεσματικότητα μια βάση δεδομένων, δύο χαρακτηριστικά που δεν υπάρχουν στον Ajax. Κρίνεται, λοιπόν, απαραίτητη η μεσολάβηση της. Όπως και πριν, έτσι και εδώ τα αποτελέσματα επιστρέφονται σε μορφή XML, τα οποία επεξεργάζονται και στην συνέχεια απεικονίζονται στην ιστοσελίδα με μετατροπή του αντίστοιχου HTML κώδικα.

Η απεικόνιση των γεωγραφικών πληροφοριών όπως είπαμε και προηγουμένως γίνεται με την υπογράμμιση τους στο κείμενο και την δημιουργία hover-over menus για κάθε μια. Για τον σκοπό αυτό χρησιμοποιήθηκε η τεχνολογία DOM, η οποία μπορεί να μετατρέψει δυναμικά τον HTML κώδικα μια ιστοσελίδας πολύ εύκολα.

Στα hover-over menus, ο χρήστης μπορεί να εκτελέσει μια από τις τρεις εναλλακτικές ενέργειες που εμφανίζονται μπροστά του, των οποίων οι παράμετροι εξαρτώνται από την γεωγραφική πληροφορία στην οποία ανήκει το κάθε menu. Με την εκτέλεση της πρώτης, “view in Google Map”, ανοίγει ένα νέο παράθυρο που περιέχει έναν Google Map με την αντίστοιχη γεωγραφική πληροφορία σημειωμένη. Η δεύτερη, “change info”, είναι ίδια με την πρώτη μόνο που στην περίπτωση που οι συντεταγμένες είναι λανθασμένες, ο χρήστης μπορεί να μετακινήσει το “icon” και να το τοποθετήσει στο σημείο που πιστεύει ότι είναι σωστότερο να τοποθετηθεί και πατώντας “OK” γίνεται η αλλαγή στην βάση. Και τέλος η τρίτη, “delete”, απλά διαγράφει την συγκεκριμένη πληροφορία από την βάση δεδομένων. Οι δύο πρώτες ανοίγουν ένα νέο παράθυρο που περιέχει έναν Google Map για την συγκεκριμένη γεωγραφική πληροφορία. Οι χάρτες αυτοί φυλάγονται στον server (Apache Tomcat). Περισσότερες πληροφορίες για την τεχνολογία των Google maps και την υλοποίησή τους στην δική μας περίπτωση υπάρχουν στις απόμενες παραγράφους. Στην τρίτη περίπτωση, όπου ζητείτε η διαγραφή ενός στοιχείου, υπάρχει κλήση προς την κεντρική βάση δεδομένων

και όπως και πριν χρησιμοποιείται το αντικείμενο XmlHttp του Ajax και με συνθετικό κρίκο την PHP.

Αφού έχει εκτελεστεί η ενέργεια της απεικόνισης της γεωγραφικής πληροφορίας, αν ο χρήστης εντοπίσει κάποια λέξη ή φράση που ενώ αποτελεί γεωγραφική πληροφορία δεν έχει εντοπιστεί από την παραπάνω διαδικασία, έχει το δικαίωμα να την προσθέσει. Ο χρήστης μαρκάρει την λέξη ή φράση με την βοήθεια του κέρσορα και εκτελεί την εντολή “geoTag” που βρίσκεται στο right-click menu του browser (Σχήμα 4.15). Στον χάρτη που εμφανίζεται ο χρήστης βρίσκει το σημείο στο οποίο αντιστοιχεί σε αυτή την λέξη ή φράση και τα δεδομένα αποθηκεύονται στην κεντρική βάση δεδομένων. Ο χάρτης που ανοίγει, όπως και προηγουμένως φυλάσσεται στον server και η ενημέρωση της βάσης γίνεται μέσα από την διαδικασία που δημιουργεί τους χάρτες (παρ 5.....).



Σχήμα 4.15 - Προσθήκη της λέξης Εύηνου

4.3.2 Java Servlet

Το κομμάτι αυτό αποτελείται από το εργαλείο αυτόματης γεωκωδικοποίησης που έχει δημιουργηθεί στην 8. Το εργαλείο αυτό δέχεται σαν είσοδο έναν πλήθος από ιστοσελίδες εντοπίζει τις γεωγραφικές πληροφορίες που περιέχουν και στην συνέχεια βρίσκει την γεωγραφική θέση της κάθε ιστοσελίδας.

Επειδή η γεωγραφική θέση της ιστοσελίδας δεν είναι ιδιαίτερα χρήσιμη για την δική μας εφαρμογή, αφού μας ενδιαφέρει εξ' ολοκλήρου ο εντοπισμός των γεωγραφικών πληροφοριών, έγινε μια μετατροπή του κώδικα του για να εφαρμοστεί στις δικές μας ανάγκες. Ο αυτόματος γεωκωδικοποιητής είναι γραμμένος στην γλώσσα προγραμματισμού Java, εκμεταλλευόμενοι αυτό το γεγονός του δίνουμε την μορφή ενός java Servlet, κάτι που κάνει την επικοινωνία του αυτόματου με το extension ιδιαίτερα εύκολη.

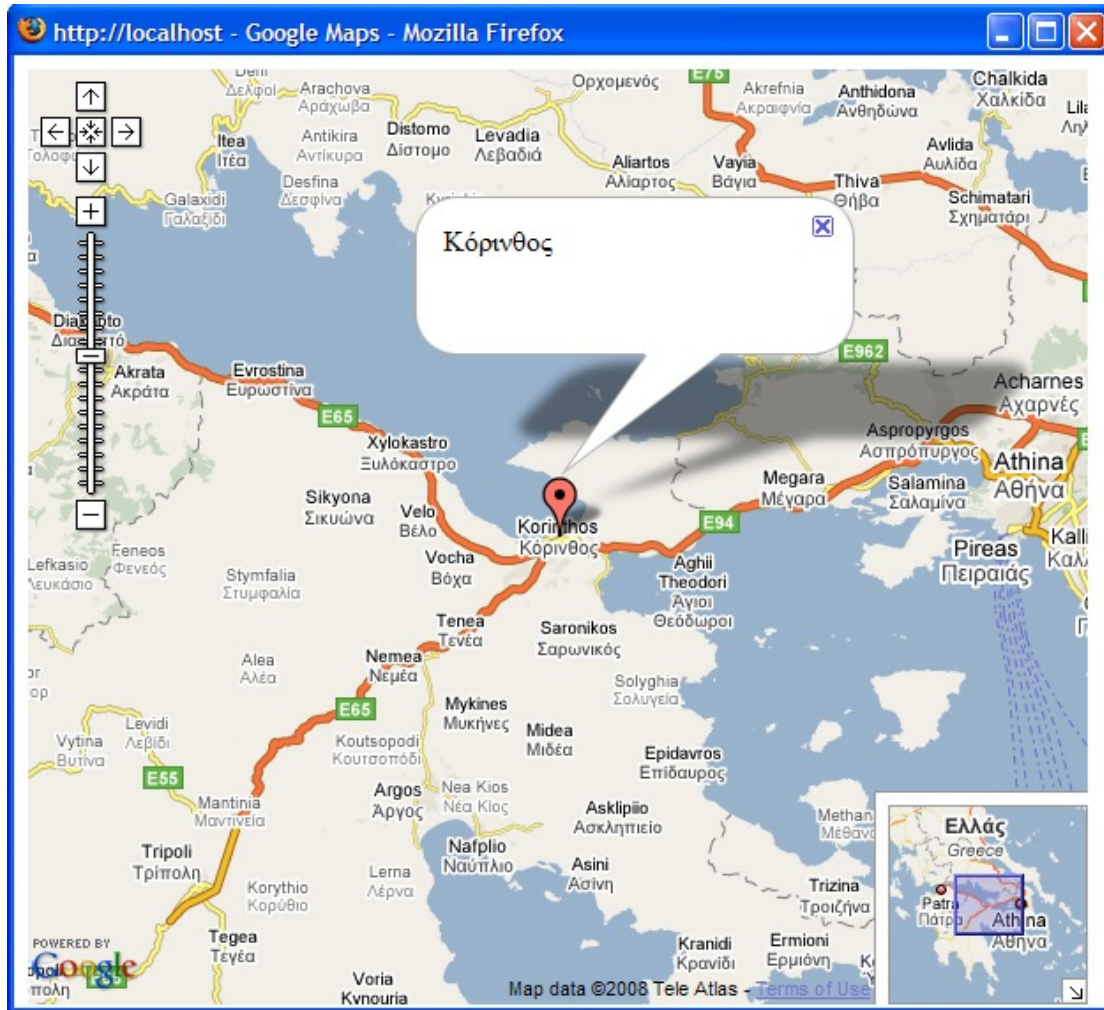
Δημιουργήθηκε μια καινούργια κλάση η *xmlHttpRequestServlet*, που ουσιαστικά αντικαθιστά την υπάρχουσα κλάση *fileCrawler*. Η *xmlHttpRequestServlet* είναι μια επέκταση της abstract κλάσης της Java *HttpServlet*. Ένα *HttpServlet* είναι μια βολική abstract

κλάση για την δημιουργία ενός servlet που δέχεται HTTP requests και παράγει HTTP responses. Επειδή στην παραπάνω κλάση οι HTTP αιτήσεις πραγματοποιούνται με την αίτηση HTTP Get, η βασική μέθοδο της είναι η doGet που δέχεται σαν παράμετρο το URL της εκάστοτε ιστοσελίδας και επιστρέφει ένα XML που περιέχει όλες τις γεωγραφικές οντότητες, που έχει πάρει σαν αποτελέσματα από την εκτέλεση του τροποποιημένου αλγόριθμου κ. Αντζελ, μαζί με σχετικές πληροφορίες για τον εντοπισμό τους στην σχετική ιστοσελίδα. Η δημιουργία αυτού του XML αρχείου γίνεται με την βοήθεια της κλάσης xmlReport (μέθοδος xmlWriter) και έχει τοποθετηθεί στο πακέτο geo (περιέχει όλες τις κλάσεις για την διαδικασία της γεωκωδικοποίησης). Περισσότερες πληροφορίες για το πως λειτουργεί ένα servlet βρίσκονται στην παράγραφο 5.5.7.

4.3.3 Visualization - Google Maps

Μέρος της εφαρμογής είναι η απεικόνιση των γεωγραφικών πληροφοριών σε γεωγραφικούς χάρτες και η παρουσίαση τους στον χρήστη. Για τον σκοπό αυτό, χρησιμοποιήθηκαν οι χάρτες της εταιρίας Google(βλ. παρ. 5.3). Η εμφάνιση αυτών των χαρτών στον χρήστη γίνεται για τους παρακάτω διαφορετικούς λόγους:

1. την παρουσίαση της γεωγραφικής πληροφορίας
2. την παρουσίαση της γεωγραφικής πληροφορίας και την αλλαγή των συντεταγμένων της από τον χρήστη, στην περίπτωση που έχουν απεικονιστεί λανθασμένα.
3. για την προσθήκη γεωγραφικής πληροφορίας που δεν έχει βρεθεί από την διαδικασία που περιγράψαμε προηγουμένως για την απεικόνιση των γεωγραφικών πληροφοριών σε μια ιστοσελίδα.



Σχήμα 4.16 – Ένας χάρτης Google

Για κάθε ένα από τα παραπάνω έχει αναπτυχθεί μια ξεχωριστή ιστοσελίδα, με βάση τα αντικείμενα και τις μεθόδους του Google Map API, και φιλοξενούνται από τον Apache Tomcat. Στο Σχήμα 4.16 παρουσιάζεται ενδεικτικά ένας τέτοιος χάρτης.

4.3.4 Persistent Storage - Βάση Δεδομένων

Για την υλοποίηση της εφαρμογής θεωρήθηκε σκόπιμη η δημιουργία ενός κεντρικού repository όπου θα είναι αποθηκευμένες όλες οι γεωγραφικές πληροφορίες από τις ιστοσελίδες στις οποίες έχει εκτελεστεί η εφαρμογή καθώς και πρόσθετες πληροφορίες για τον εντοπισμό τους σε αυτές. Λόγω της απλότητας της εφαρμογής, για την υλοποίηση της βάσης δεδομένων χρησιμοποιήθηκε ένας απλός πίνακας, αφού οποιοδήποτε άλλο σχεσιακό μοντέλο θα ήταν απλά περιττό.

5

Λεπτομέρειες

Υλοποίησης

Στην συνέχεια ακολουθεί μια περιγραφή των εργαλείων και των τεχνολογιών που χρησιμοποιήθηκαν.

5.1 *Mozilla FireFox*

Ο **Mozilla FireFox** 88(γνωστός και ως **FireFox**, αρχικά ως *Phoenix* και για λίγο ως *Mozilla Firebird*) είναι ένας ελεύθερος, γραφικός φυλλομετρητής (Web browser) σε πολλά λειτουργικά συστήματα, που αναπτύχθηκε από το Mozilla Foundation και εκατοντάδες εθελοντών. Αποτελεί έναν από του ασφαλέστερους τρόπους για περιήγηση στο Ιστό. Ο FireFox είναι βασισμένος σε κώδικα ενός άλλου προϊόντος του Mozilla Foundation, του Navigator που είναι ο φυλλομετρητής - μέρος του Mozilla Application Suite. Η παράλληλη δημιουργία του FireFox προσανατολίστηκε στην ανάπτυξη ενός ελαφρού, γρήγορου και εύκολα επεκτάσιμου φυλλομετρητή, και τελικά αντικατέστησε (μαζί με τον Thunderbird ως εφαρμογή e-mail) τον κύριο αναπτυξιακό προσανατολισμό του Mozilla Foundation. Για την υλοποίηση χρησιμοποιήθηκε ο Mozilla FireFox 2.0.0.10.

5.2 *Mozilla extension*

Τα **Extensions** δείναι εγκαταστήσιμα συμπληρωματικά προγράμματα του Mozilla Foundation's projects (επίσης είναι συμβατά με διάφορες παραλλαγές του όπως το Portable FireFox), τα οποία μπορούν να προσθέσουν νέα χαρακτηριστικά σε μία εφαρμογή ή ακόμα και να τροποποιήσουν τα ήδη υπάρχοντα. Οι σχεδιαστές του Mozilla ήθελαν ο FireFox να

είναι μια minimal εφαρμογή, ώστε να αποφευχθούν τα software bloats και τα bugs, που ο κάθε χρήστης θα μπορούσε να διαμορφώσει τον δικό του browser με τα χαρακτηριστικά που ο ίδιο προτιμάει. Για την δημιουργία ενός extension χρησιμοποιούνται οι ακόλουθες τεχνολογίες :

- **XPI** - Cross-Platform Installer module
- **JavaScript** (βλ. παρ) – είναι η κύρια γλώσσα για την δημιουργία extensions
- **XUL** 888(**XML** User Interface Language) – χρησιμοποιείται για να προσδιορίσει του **UI** (User Interface) και την αλληλεπίδραση του extension με τους χρήστες.
- **DOM** 88(Document Object Model) – Μπορεί να αλλάξει δυναμικά το XUL ή ακόμα και να μετατρέψει ένα **HTML** που μόλις έχει φορτωθεί.
- **CSS** (Cascading Style Sheets) γλώσσα που περιγράφει την εμφάνιση ενός HTML εγγράφου(**style sheet language**)
- **XPCOM/XPCConnect**

Τα extensions χρησιμοποιούνται κυρίως για να προσθέσουν πρόσθετες λειτουργίες σε έναν φυλλομετρητή (Browser). Παραδείγματα τέτοιων λειτουργιών είναι οι αναγνώστες RSS, bookmark organizers, toolbars, website-specific client programs, FTP, e-mail, mouse gestures, proxy server switching, or debugging tools.

Πολλά extensions μπορούν να αλλάξουν το περιεχόμενο μιας ιστοσελίδας , όχι την ίδια την ιστοσελίδα , αλλά τον τρόπο με τον οποίο εμφανίζεται στον Browser ενός μεμονωμένου χρήστη. Για παράδειγμα, τα extensions Adblock μπορούν να αποτρέψουν τον browser από τις εικόνες φόρτωσης που είναι διαφημίσεις. Ένα άλλο δημοφιλές extension το CustomizeGoogle, αλλάζει τον τρόπο που οι σελίδες Google εμφανίζονται, με την παροχή των συνδέσεων με άλλες μηχανές αναζήτησης. Επίσης, το extension Greasemonkey επιτρέπει στο χρήστη για να εγκαταστήσει τα scripts που τροποποιούν την εμφάνιση των συγκεκριμένων σελίδων ή του Ιστού γενικά.

Τέλος, κάποια extensions μπορεί να έχουν σατιρικό ή χιουμοριστικό περιεχόμενο. .

5.2.1 *To Hello, World! του extension*

Το extension που θα κατασκευάσουμε για αυτό το παράδειγμα είναι αρκετά απλό προσθέτει την εντολή “Hello, World!” στο μενού Εργαλείων (Tools Menu) του φυλλομετρητή, που όταν εκτελείται θα εμφανίζει ένα παράθυρο που περιέχει την φράση “Hello, World!” 888 .

5.2.1.1 Σχεδίαση του Extension

Αρχικά θα πρέπει να σχεδιάσουμε το πώς θα είναι το extension. Τα ακόλουθα αρχεία είναι όλα απαραίτητα για την εφαρμογή μας (μπορούμε να αλλάξουμε τα ονόματα σε όλα τα αρχεία εκτός από τα ονόματα των φακέλων και των αρχείων “chrome.manifest” και “install.rdf”).

Όνομα Αρχείου	Περιγραφή
chrome.manifest	Ενημερώνει τον Firefox που είναι τα αρχεία “chrome” πώς να τα αξιοποιήσει
install.rdf	Αρχείο που περιγράφει το extension (“Install manifest”)
overlay.xul	Περιγράφει τα στοιχεία του περιβάλλοντος χρήστη που προστίθενται στο παράθυρο του φυλλομετρητή.
overlay.js	Περιέχει τον κώδικα που εκτελείται στο παράθυρο του φυλλομετρητή.
overlay.dtd	Περιέχει μεταφράσεις διάφορων εκφράσεων που υπάρχουν στο overlay.xul.
hello.dtd	Περιέχει μεταφράσεις διάφορων εκφράσεων που υπάρχουν στο hello.xul.
overlay.css	Ρυθμίζει την εμφάνιση του περιβάλλοντος χρήστη χρησιμοποιώντας την τεχνολογία CSS
hello.xul	Επιγράφει τα στοιχεία του περιβάλλοντος χρήστη στο νέο παράθυρο που δημιουργείται
helloworld@mozilla.doslash.org	Ένας δείκτης στα αρχεία του extension

Στην συνέχεια φαίνεται η δομή που πρέπει να έχει το extension και το πώς τοποθετούνται τα προηγούμενα αρχεία, έτσι ώστε ο φυλλομετρητής Mozilla FireFox να τα αντιληφθεί.

helloworld/ chrome.manifest install.rdf content/ overlay.js overlay.xul hello.xul locale/ en-US/ overlay.dtd hello.dtd skin/ overlay.css
--

Πίνακας 1 – Δομή του extension “Hello, World!”

Η δομή αυτή μπορεί να αλλαχθεί μόνον αλλάξουμε τα περιεχόμενα του chrome.manifest, αλλά για λόγους ομοιογένειας προτιμάται να μείνει ως έχει.

5.2.1.2 Επεξήγηση Αρχείων

Στο σημείο αυτό γίνεται μια επεξήγηση της δομής και του σκοπού των αρχείων του extension “Hello, World!”.

- **Chrome.manifest**

Οι πιο πρόσφατες εκδόσεις του Firefox διαβάζουν ένα απλό κείμενο το chrome.manifest 8 για να δουν τι αρχεία περιέχει ένα extension. Στο παράδειγμα που ακολουθεί δημιουργούμενα ένα αρχείο τύπου chrome.manifest για το extension “Hello, world!”(Θεωρούμε ότι το όνομα του extension είναι “helloworld”).

```
content helloworld content/  
overlay chrome://browser/content/browser.xul chrome://helloworld/content/overlay.xul  
  
locale helloworld en-US locale/en-US/  
  
skin helloworld classic/1.0 skin/  
style chrome://global/content/customizeToolbar.xul chrome://helloworld/skin/overlay.css
```

Πίνακας 2 – Αρχείο chrome.manifest

Στην συνέχεια αναλύουμε τι κάνει κάθε σειρά :

Η σειρά 1 πληροφορεί για το που βρίσκονται τα αρχεία του extension που προσφέρουν τις καινούργιες ιδιότητες στον φυλλομετρητή που στην δική μας περίπτωση είναι στο φάκελο με το όνομα content.

Η σειρά 2 επιτρέπει στο overlay.xul να τροποποιήσει το περιβάλλον χρήστη στο κεντρικό παράθυρο του FireFox.

Η σειρά 4 ενημερώνει που βρίσκονται οι en-US φράσεις.

Η σειρά 6 ενημερώνει για το που βρίσκονται οι εικόνες που θα χρησιμοποιηθούν.

Ενώ η σειρά 7 μας πληροφορεί ότι το overlay.cuss πάνω θα εφαρμοστεί στον έγγραφο chrome://global/content/customizeToolbar.xul του FireFox.

Σημείωση : Αυτό το αρχείο αλλάζει κάπως όταν θέλουμε να δημιουργήσουμε ένα αρχείο XPI για το extension μας.

- **Overlay.xul**

Το αρχείο overlay.xul 88 είναι ένα απλό αρχείο τύπου XUL που χρησιμοποιείται για επικάλυψη κάποιων ιδιοτήτων του φυλλομετρητή με καινούργιες. Το overlay που χρησιμοποιήσαμε στο παράδειγμά μας είναι το ακόλουθο :


```

<?xml version="1.0"?>
<?xml-stylesheet href="chrome://helloworld/skin/overlay.css"
type="text/css"?>
<!DOCTYPE overlay SYSTEM
"chrome://helloworld/locale/overlay.dtd">
<overlay id="helloworld-overlay"
xmlns="http://www.mozilla.org/keymaster/gatekeeper/there.is.only.xul">
  <script src="overlay.js"/>
  <menupopup id="menu_ToolsPopup">
    <menuitem id="helloworld-hello" label="&helloworld;"
      oncommand="Helloworld.onMenuItemCommand(event);"/>
  </menupopup>
</overlay>

```

Πίνακας 3 – Overlay.xul

Ακολουθεί ανάλυση με το τι κάνει κάθε σειρά :

Σειρά 1 : Γνωστοποιεί πια έκδοση του XML χρησιμοποιείται.

Σειρά 2 : Είναι προαιρετική και χρησιμοποιείται για να εφαρμόσει ένα έγγραφο με στυλ στο βασικό έγγραφο

Σειρά 3 : Είναι επίσης προαιρετική, αλλά συνίσταται η χρήση της αν περιέχονται διάφορες εκφράσεις ορατές στους χρήστες, αφού βοηθάει στην τοπικότητα του extension.

Οι σειρές 4-5 καθορίζουν το κεντρικό στοιχείο του overlay. Πρέπει να είναι ένα στοιχείο XUL <overlay>.

Η σειρά 6 είναι προαιρετική προσθέτει αρχεία γραμμένα σε γλώσσα JavaScript στο κυρίως έγγραφο.

Τέλος οι σειρές 8-10 προσθέτουν ένα νέο αντικείμενο στο Μενού εργαλείων του κεντρικού παραθύρου του Firefox.

- **Overlay.js**

Το αρχείο αυτό περιέχει κώδικα γραμμένο σε JavaScript που καθορίζει τον τρόπο συμπεριφοράς της εφαρμογής μας. Ο κώδικας που χρησιμοποιήθηκε για το “Hello, World!” είναι ο ακόλουθος :

```

var Helloworld = {
  onLoad: function() {
    // initialization code
    this.initialized = true;
  },

  onMenuItemCommand: function() {
    window.open("chrome://helloworld/content/hello.xul", "", "chrome");
  }
};

```

```
}  
};  
window.addEventListener("load", function(e) { Helloworld.onLoad(e); }, false);
```

Πίνακας 4 – Overlay.js

- **Overlay.dtd**

Τα αρχεία DTD χρησιμοποιούνται για να μετατρέψουν τα αρχεία XUL/XBL/XHTML και άλλα αρχεία XML του Mozilla τοπικά. Βασικά, αντί να γράφουμε τις διάφορες εκφράσεις κατευθείαν στο αρχείο XUL, χρησιμοποιούμε τις οντότητες XML (**XML entities**), οι οποίες παίρνουν τις τιμές του, που τους έχουν δοθεί από το αντίστοιχο αρχείο DTD.

Αυτή η διαδικασία προσφέρει στο extension μια τοπικότητα, επειδή μπορεί να υπάρχουν πολλά διαφορετικά τοπικά σύνολο και ο Mozilla θα διαλέξει κάποιο από αυτά την ώρα της εκτέλεσης της εφαρμογής.

Στο δικό μας παράδειγμα απλά για να δείξουμε πως λειτουργεί προσθέσαμε τον ακόλουθο κωδικό :

```
<!ENTITY helloworld "Hello World!">
```

Πίνακας 5 – Overlay.dtd

- hello.dtd

Αυτό το αρχείο είναι για το καινούργιο παράθυρο

```
<!ENTITY title.label "Hello World">  
<!ENTITY separate.label "This is a separate window!">  
<!ENTITY close.label "Close">
```

Πίνακας 6 – Hello.dtd

- hello.xul

Εκτός από το να τροποποιήσουμε τα υπάρχοντα παράθυρα, μπορούμε να δημιουργήσουμε καινούργια. Το περιβάλλον χρήστη των παραθύρων έχει κατασκευαστεί από αρχεία XUL, αλλά αντίθετα από τα overlays, το αρχικό στοιχείο είναι είτε το <window> είτε το <dialog> και όχι το <overlay>. Μια άλλη διαφορά είναι ότι δεν χρειάζεται να δηλώσουμε στο chrome.manifest το κάθε νέο παράθυρο, όπως στην περίπτωση του overlay. Το παράδειγμα μας περιέχει το αρχείο hello.xul, που χρησιμοποιείται για την δημιουργία ενός νέου παραθύρου.

```
<?xml version="1.0"?>  
<?xml-stylesheet href="chrome://global/skin/global.css" type="text/css"?>  
<!DOCTYPE window SYSTEM "chrome://helloworld/locale/hello.dtd">  
  
<window xmlns="http://www.mozilla.org/keymaster/gatekeeper/there.is.only.xul"  
title="&title.label;">
```

```

<hbox align="center">
  <description flex="1">&separate.label;</description>
  <button label="&close.label;" oncommand="close();"/>
</hbox>

</window>

```

Πίνακας 7 – Hello.xul

- **install.rdf**

Το `install.rdf` 8, όπως μπορούμε εύκολα να διαπιστώσουμε, είναι ένα αρχείο το οποίο περιέχει γενικές πληροφορίες για το `extension` όπως το ID του, το όνομα, μια περιγραφή του, το όνομα του συγγραφέα του κ.τ.λ. Επίσης, το `Target Application` περιέχει πληροφορίες για την εφαρμογή για την οποία είναι φτιαγμένο. Έτσι, το `<em:id>` με το νούμερο `EC8030F7...` Δηλώνει ότι το `extension` αυτό ενσωματώνεται στον Mozilla Firefox και τα `minVersion` και `maxVersion`, την ελάχιστη και μέγιστη απαιτούμενη έκδοση του φυλλομετρητή. Για να είναι συμβατό το `extension` με τις επόμενες εκδόσεις του Mozilla Firefox, δεδομένου ότι βελτιώνεται αρκετά συχνά, σαν `maxVersion` βάζουμε την 9.9 ώστε να μην ανατρέχουμε συνέχεια στο `extension` μας και να ανανεώνουμε την `maxVersion`.

```

<?xml version="1.0"?>
<RDF xmlns="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:em="http://www.mozilla.org/2004/em-rdf#">

  <Description about="urn:mozilla:install-manifest">

    <em:id>helloworld@mozilla.doslash.org</em:id>
    <em:name>Hello World (Firefox 1.5 or later edition)</em:name>
    <em:version>1.0</em:version>
    <em:description>Classic first extension from MozillaZine KB</em:description>
    <em:creator>Nickolay Ponomarev</em:creator>
    <!-- optional items -->
    <em:contributor>A person who helped you</em:contributor>
    <em:contributor>Another one</em:contributor>

    <em:homepageURL>http://kb.mozillazine.org/Getting_started_with_extension_development</em:homepageURL>
    <!-- em:optionsURL>chrome://sampleext/content/settings.xul</em:optionsURL>
    <em:aboutURL>chrome://sampleext/content/about.xul</em:aboutURL>
    <em:iconURL>chrome://sampleext/skin/mainicon.png</em:iconURL>

    <em:updateURL>http://sampleextension.mozdev.org/update.rdf</em:updateURL-->

    <!-- Firefox -->
    <em:targetApplication>
      <Description>

```

```
<em:id>{ec8030f7-c20a-464f-9b0e-13a3a9e97384}</em:id>
<em:minVersion>1.5</em:minVersion>
<em:maxVersion>9.9</em:maxVersion>
</Description>
</em:targetApplication>

</Description>

</RDF>
```

Πίνακας 8 – Install.rdf

- overlay.css

Το overlay.css χρησιμοποιείται από τον Firefox για να εφαρμόσει μια συγκεκριμένη εμφάνιση στα στοιχεία του extension. Για το παράδειγμά μας, χρωματίζουμε το “helloworld” που τοποθετήσαμε στο Μενού Εργαλείων με κόκκινο.

Σημείωση : Ο χρωματισμός των αντικειμένων των μενού δεν συνίσταται.

```
menuitem#helloworld-hello {
  color: red !important;
}
```

Πίνακας 9 – Overlay.css

5.2.1.3 Αναγνώριση του Extension από τον Διαχειριστή των Extensions

Πριν εκδοθεί το extension πρέπει να δοκιμαστεί. Για τα αποφευχθούν τυχόν προβλήματα κατά την διάρκεια των δοκιμών, δημιουργούμε και χρησιμοποιούμε ένα νέο προφίλ 8 στον Firefox. Αφού έχουμε δημιουργήσει τα απαραίτητα αρχεία και τα έχουμε τοποθετήσει με την δομή που περιγράψαμε προηγουμένως, εγγράφουμε το extension στον διαχειριστή extensions (Extension Manager) δημιουργώντας απλά ένα αρχείο «δείκτη» με το όνομα του extension ID. Το αρχείο αυτό περιέχει το μονοπάτι για τους φακέλους install.rdf και chrome.manifest.

Στο παράδειγμα μας το ID είναι το helloworld και ο δείκτης ο helloworld@mozilla.doslash.org και θέλουμε να γίνει εγγραφή του X:\Dev\helloworld\ . Έτσι το αρχείο helloworld@mozilla.doslash.org θα περιέχει τα ακόλουθα :

```
X:\Dev\helloworld
```

Πίνακας 10 – Extension Pointer

Στην συνέχεια επανεκκινούμε τον Firefox και ελέγχουμε αν πραγματοποιήθηκε η εγκατάσταση του extension.

5.2.1.4 «Πακετάρισμα»

Όπως προαναφέρθηκε, η δομή των φακέλων του extension και το αρχείου chrome.manifest πρέπει να αλλάξουν πριν γίνει το «πακετάρισμα» (Packaging).

Ακολουθούν οι δύο διαφορετικές δομές :

Developing	Packaging
Δομή Φακέλου – Folder Structure	
helloworld/ chrome.manifest install.rdf components/ defaults/ preferences/ mydefaults.js content/ overlay.js overlay.xul locale/ en-US/ overlay.dtd skin/ overlay.css	helloworld.xpi/ chrome.manifest install.rdf components/ defaults/ preferences/ mydefaults.js chrome/ helloworld.jar content/ overlay.js overlay.xul locale/ en-US/ overlay.dtd skin/ overlay.css

Πίνακας 11 – Δομή του extension στις περιπτώσεις Developing και Packaging

chrome.manifest – Developing		
content	helloworld content/	
overlay	chrome://browser/content/browser.xul	chrome://helloworld/content/overlay.xul
locale	helloworld en-US	locale/en-US/
skin	helloworld classic/1.0 skin/	
style	chrome://global/content/customizeToolbar.xul	chrome://helloworld/skin/overlay.css

Πίνακας 12 – Αρχείο chrome.manifest στην περίπτωση Developing

chrome.manifest – Packaging (Προσοχή στο «jar:chrome/helloworld.jar!/»)		
content	helloworld jar:chrome/helloworld.jar!/content/	
overlay	chrome://browser/content/browser.xul	chrome://helloworld/content/overlay.xul
locale	helloworld en-US	jar:chrome/helloworld.jar!/locale/en-US/
skin	helloworld classic/1.0 jar:chrome/helloworld.jar!/skin/	

style	chrome://global/content/customizeToolbar.xul	chrome://helloworld/skin/over-
lay.css		

Πίνακας 13 - Αρχείο chrome.manifest στην περίπτωση Packaging

5.2.2 Το δικό μας Extension

Στην παράγραφο 5.2.1.1 περιγράψαμε την δομή που έχει το extension “Hello, World!”, μία δομή που πρέπει να έχει κάθε extension. Έτσι λοιπόν, το extension που χρησιμοποιήσαμε για την υλοποίηση της εφαρμογής μας έχει μια αντίστοιχη δομή η οποία περιγράφεται από τον παρακάτω Πίνακας 14 :

```
geoJoy/  
  chrome.manifest  
  install.rdf  
  content/  
    overlay.js  
    overlay.xul  
  locale/  
    en-US/  
  skin/
```

Πίνακας 14-Δομή extension GeoJoy

Όπως μπορεί εύκολα να παρατηρήσει κανείς οι φάκελοι locale και skin δεν περιέχουν αρχεία αφού αρχεία της μορφής CSS και DTD να προστεθούν στην εφαρμογή μας.

Στην συνέχεια ακολουθεί μία περιγραφή των αρχείων που χρησιμοποιήθηκαν.

- **Chrome.manifest**

```
content geojoy content/  
overlay chrome://browser/content/browser.xul chrome://geojoy/content/overlay.xul
```

Πίνακας 15 – GeoJoy : chrome.manifest

Βλέπουμε ότι στο chrome.manifest δηλώνονται μόνο το αρχείο “overlay.xul” αφού αρχεία CSS και DTD δεν χρησιμοποιήθηκαν (όπως στο “Hello, World!”).

- **Install.rdf**

```
<?xml version="1.0"?>  
<RDF xmlns="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:em="http://www.mozilla.org/2004/em-rdf#">  
  <Description about="urn:mozilla:install-manifest">  
    <em:id>geojoy@mozilla.doslash.org</em:id>  
    <em:name>geojoy</em:name>  
    <em:version>1.0</em:version>  
    <em:type>2</em:type>  
    <em:description>An extension for Geotagging Web Pages</em:description>  
    <em:creator>Lontou Charikleia</em:creator>  
    <!-- optional items -->  
    <em:contributor>Dieter Pfoser</em:contributor>  
    <em:homepageURL>http://kb.mozillazine.org/Getting_started_with_extension_development</em:homepageURL>
```

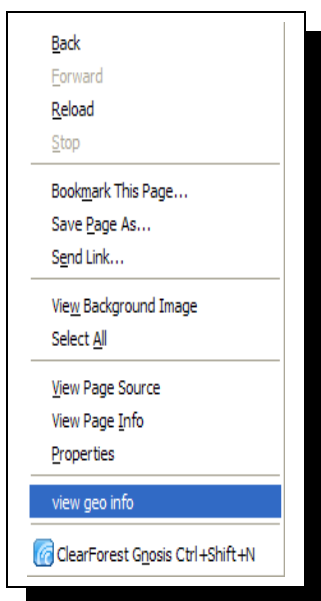
```
<!-- em:optionsURL>chrome://sampleext/content/settings.xul</em:optionsURL>
<em:aboutURL>chrome://sampleext/content/about.xul</em:aboutURL>
<em:iconURL>chrome://sampleext/skin/mainicon.png</em:iconURL>
em:updateURL>http://sampleextension.mozdev.org/update.rdf</em:updateURL-->
<!-- Firefox -->
<em:targetApplication>
  <Description>
    <em:id>{ec8030f7-c20a-464f-9b0e-13a3a9e97384}</em:id>
    <em:minVersion>1.5</em:minVersion>
    <em:maxVersion>2.0.0.*</em:maxVersion>
  </Description>
</em:targetApplication>
</Description>
</RDF>
```

Πίνακας 16 - GeoJoy : install.rdf

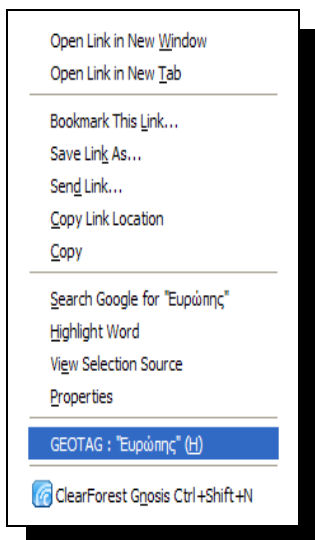
Ξέρουμε πλέον ότι το αρχείο install.rdf περιέχει γενικές πληροφορίες για το extension. Έτσι στον Πίνακα 16 παρατηρούμε τα στοιχεία του extension της εφαρμογής μας.

- **Overlay.xul**

Τα αρχεία XUL χρησιμοποιούνται για την επικάλυψη κάποιων ιδιοτήτων του φυλλομετρητή με νέες. Επίσης, χρησιμοποιούνται για να προσθέσουν και καινούργιες ιδιότητες, όπως κάνει και το αρχείο “overlay.xul” που δημιουργήσαμε. Το “overlay.xul” προσθέτει στο “right-click menu” του φυλλομετρητή τις εντολές “view geo info” και “geotag : ...”. Οι εντολές αυτές δεν εμφανίζονται ποτέ μαζί, αφού η “geotag : ...” εμφανίζεται μόνο όταν έχουμε μαρκάρει κάποιο κείμενο (Σχήμα 5.18), ενώ η “view geo info” (Σχήμα 5.17) σε όλες τις άλλες περιπτώσεις.



Σχήμα 5.17 – Right-click menu : view geo info



Σχήμα 5.18 – Right-click menu : geotag “Ευρώπης”

Στην συνέχεια παραθέτουμε τον κώδικα που χρησιμοποιήθηκε :

```
<?xml version="1.0"?>
<?xml-stylesheet href="chrome://global/skin/xul.css" type="text/css"?>
<!DOCTYPE overlay>
<overlay id="geojoy-overlay"
  xmlns:html="http://www.w3.org/1999/xhtml"
  xmlns="http://www.mozilla.org/keymaster/gatekeeper/there.is.only.xul"
  >
  <html:script language="javascript"
src="chrome://geojoy/content/overlay.js"/>
  <popup id="contentAreaContextMenu">
    <menuseparator/>
    <menuitem id="geo_tag" accesskey="H"
oncommand="GeoTag.onMenuItemCommand(event);"/>
    <menuitem id="geo_info" label = "view geo info"
oncommand = "GeoTag.view_geo_info(event);"/>
  </popup>
</overlay>
```

Πίνακας 17 - GeoJoy : Overlay.xul

- **Overlay.js**

Το αρχείο αυτό περιέχει τον κώδικα στην γλώσσα προγραμματισμού JavaScript που απαιτείται. Αποτελείται από συναρτήσεις που φορτώνουν και αρχικοποιούν το extension, ρυθμίζουν την επικοινωνία με τον κεντρικό εξυπηρετητή έτσι ώστε να εμφανιστούν οι χάρτες, να τεθεί σε λειτουργία το Java Servlet και να επεξεργαστεί η κεντρική βάση δεδομένων, βρίσκουν τις γεωγραφικές οντότητες και τις υπογραμμίζουν και τέλος δημιουργούν τα “hover-over menus” (Σχήμα 4.14).

5.3 Google MAPS API

Η Google δημιούργησε τους Google χάρτες API για να διευκολύνει την ενσωμάτωση των χαρτών Google στους ιστοχώρους από των καθένα και αποτελεί ελεύθερη υπηρεσία. Με τη

χρησιμοποίηση των Google χαρτών API μπορούμε να ενσωματώσουμε τους πλήρεις χάρτες Google σε έναν εξωτερικό ιστοχώρο. Με την προσθήκη κώδικα Google JavaScript σε μια ιστοσελίδα μπορούμε να δημιουργήσουμε το δικό μας Google maps προσαρμοσμένο στις ανάγκες μας.

Μπορεί κανείς με τις λειτουργίες που παρέχονται να σχεδιάσει δείκτες πάνω στο χάρτη, ή να αναπτύξει ακόμα πιο πολύπλοκες εφαρμογές. Προς το παρόν όμως οι υπηρεσίες αυτές που παρέχονται από την Google είναι διαθέσιμες μόνο για Web σελίδες και δεν μπορούν να χρησιμοποιηθούν από άλλη εφαρμογή. Δε βασίζονται σε κάποιο ανοιχτό πρότυπο, όπως το SOAP/XML αλλά χρησιμοποιεί JavaScript όπως προαναφέραμε. Για το λόγο αυτό, ο μόνος τρόπος ενσωμάτωσης τους είναι σε ιστοσελίδες.

Το πλεονέκτημα από τη χρήση των υπηρεσιών αυτών είναι ότι παρέχονται δωρεάν (τουλάχιστον για Web Sites με μέγιστο αριθμό επισκέψεων 50000/ημέρα). Το μόνο που χρειάζεται να κάνει κάποιος για να χρησιμοποιήσει τις υπηρεσίες αυτές είναι να εγγραφεί στο σύστημα της Google για να του παραχωρηθεί ένας κωδικός πρόσβασης στα API της Google (API key). Εφόσον όμως οι υπηρεσίες παρέχονται δωρεάν από τη Google θα πρέπει όποιος τις χρησιμοποιεί να μην τις διαθέτει στο site του επί πληρωμή αλλά επίσης δωρεάν προς τους επισκέπτες του ιστοχώρου του. Όταν το API προώθησε αρχικά, δεν είχε τη δυνατότητα να γεοκωδικοποιεί διευθύνσεις, και απαιτούσε από τον χρήστη να προσθέσει με το χέρι τα σημεία (γεωγραφικό πλάτος, γεωγραφικό μήκος) το σχήμα.

5.3.1 Προσφερόμενες υπηρεσίες

Στη συνέχεια παρουσιάζουμε συνοπτικά τα χαρακτηριστικά των υπηρεσιών που παρέχονται στο Google Maps API:

- **The GMap class:** Πρόκειται για την βασική κλάση. Ένα αντικείμενο της κλάσης GMap αντιστοιχεί σε ένα χάρτη στην σελίδα μας. Μπορεί κανείς να δημιουργήσει όσα στιγμιότυπα αυτής της κλάσης επιθυμεί (ένα για κάθε χάρτη στη σελίδα). Όταν δημιουργούμε ένα νέο στιγμιότυπο χάρτη, καθορίζουμε ένα στοιχείο στη σελίδα το οποίο θα περιέχει το χάρτη. Ο χάρτης στη συνέχεια, σαν μέγεθος του χρησιμοποιεί το μέγεθος του στοιχείου που τον περιλαμβάνει εκτός και αν το ορίσουμε διαφορετικά. Η κλάση GMap παρέχει μεθόδους χειρισμού του κέντρου του χάρτη και του επιπέδου zoom, καθώς και μεθόδους για προσθήκη ή αφαίρεση διαφόρων overlays (όπως για παράδειγμα στιγμιότυπα των κλάσεων GMarker και GPolyline όπως θα δούμε). Επιπλέον, παρέχει μεθόδους που μας δίνουν τη δυνατότητα να ανοίξουμε ένα «παράθυρο πληροφοριών» το οποίο θα περιέχει διάφορες πληροφορίες πάνω στο χάρτη.
- **Events:** Με τη χρήση των event listeners μπορούμε να εισάγουμε δυναμικά στοιχεία στην

εφαρμογή μας. Ένα αντικείμενο της κλάσης αυτής παρέχει έναν αριθμό από γεγονότα (events) και η εφαρμογή μας μπορεί να τα «ακούει» χρησιμοποιώντας τις στατικές μεθόδους `GEvent.addListener` ή `GEvent.bind`. Έτσι μπορεί το πρόγραμμα μας για παράδειγμα να εμφανίζει ένα μήνυμα ανάλογα με το κλικ ενός χρήστη πάνω στο χάρτη.

- **The Info Window:** Κάθε χάρτης έχει ένα μόνο «παράθυρο πληροφοριών», το οποίο εμφανίζει HTML περιεχόμενο σε ένα παράθυρο πάνω από το χάρτη. Το παράθυρο πληροφοριών μοιάζει με ένα «συννεφάκι» σε ένα βιβλίο comic. Αποτελείται από μια περιοχή με το περιεχόμενο των πληροφοριών η οποία λεπταίνει στην άκρη και γίνεται σαν δείκτης που δείχνει σε ένα καθορισμένο σημείο πάνω στο χάρτη. Αν κάποιος έχει χρησιμοποιήσει το Google Maps ή το Google Local, τότε πολύ πιθανό να έχει δει ένα «παράθυρο πληροφοριών» όταν κάνει κλικ σε ένα εικονίδιο (marker). Άλλο χαρακτηριστικό των παραθύρων αυτών είναι ότι δεν μπορεί κανείς να εμφανίσει περισσότερα από ένα ταυτόχρονα σε ένα δοσμένο χάρτη αλλά μπορεί να μετακινήσει το παράθυρο και να αλλάξει τα περιεχόμενά του αν αυτό είναι επιθυμητό.

Η βασική μέθοδος για ένα παράθυρο πληροφοριών είναι η `openInfoWindow`, η οποία παίρνει σαν είσοδο ένα σημείο και ένα HTML DOM element. Το παράθυρο πληροφοριών εμφανίζεται με το κείμενό του στο δοσμένο σημείο του χάρτη και εμφανίζει το DOM element στην περιοχή που περιλαμβάνει. Η μέθοδος `openInfoWindowHtml` είναι παρόμοια, αλλά παίρνει ένα HTML string σαν δεύτερο όρισμα αντί για ένα DOM element. Ομοίως, η `openInfoWindowXslt` παίρνει ένα σημείο, ένα XML DOM element, και το URL ενός XSLT αρχείου. Στη συνέχεια εφαρμόζει το μετασχηματισμό XSLT στο XML για να παράγει τα περιεχόμενα του παραθύρου. Η μέθοδος αυτή μεταφέρει το XSLT ασύγχρονα αν δεν έχει ήδη μεταφερθεί από τον browser του χρήστη.

Εκτός από τα παραπάνω, μπορούμε επίσης να εμφανίσουμε ένα παράθυρο πληροφοριών πάνω από ένα overlay όπως για παράδειγμα ένα εικονίδιο (marker). Για να το κάνουμε αυτό περνάμε ως τρίτο όρισμα ένα pixel offset μεταξύ του καθορισμένου σημείου και του κειμένου του παραθύρου πληροφοριών. Η κλάση `GMarker` (την οποία θα δούμε παρακάτω) επιτρέπει μεθόδους `openInfoWindow` οι οποίες χειρίζονται τα pixel offsets αυτόματα βασισμένες στο μέγεθος και σχήμα του εικονιδίου, και συνεπώς δε χρειάζεται να ανησυχεί ο προγραμματιστής για τον υπολογισμό των offsets στην εφαρμογή του.

- **Overlays:** Τα overlays είναι αντικείμενα πάνω στο χάρτη τα οποία είναι «συνδεδεμένα» σε γεωγραφικές συντεταγμένες, και έτσι μετακινούνται όταν ο χρήστης μετακινεί το χάρτη ή κάνει zoom σε αυτόν ή όταν αλλάζει ο τρόπος προβολής (η Google παρέχει διάφορους τρόπους προβολής ενός χάρτη, για παράδειγμα δορυφορική προβολή κτλ). Το Google Maps API περιλαμβάνει δύο τύπους overlays, τα markers που είναι εικονίδια πάνω στο χάρτη και τα polylines που είναι γραμμές που κατασκευάζονται από μία σειρά

σημείων.

- **Markers and Icons:** Ο GMarker constructor παίρνει σαν είσοδο ένα εικονίδιο και ένα σημείο και παράγει ένα μικρό σύνολο από events όπως για παράδειγμα το «κλικ» το οποίο μπορούμε να χειριστούμε στη συνέχεια στον κώδικά μας. Το δυσκολότερο κομμάτι στη δημιουργία ενός marker είναι ο καθορισμός του εικονιδίου. Αυτό είναι δύσκολο διότι είναι μεγάλος ο αριθμός των διαφορετικών εικόνων οι οποίες συνθέτουν ένα απλό εικονίδιο στο Maps API. Παρόλα αυτά, αν θέλει κανείς ένα γενικό εικονίδιο μπορεί να δημιουργήσει ένα GMarker χωρίς να καθορίσει ένα εικονίδιο.

Τα εικονίδια είναι συνήθως της μορφής πινέζας, με ένα tip το οποίο εμφανίζεται στη θέση που καθορίζεται από τον GMarker constructor. Κάθε εικονίδιο έχει (τουλάχιστον) μια εικόνα στο προσκήνιο και μία εικόνα σαν σκιά. Υπάρχουν και περισσότερες λεπτομέρειες που αφορούν το πώς πρέπει να καθορίζονται τα εικονίδια και μπορούν να αναζητηθούν στο documentation του Google Maps API .

- **Polylines:** Ο GPolyline constructor παίρνει σαν είσοδο έναν πίνακα σημείων και δημιουργεί μια σειρά από ευθύγραμμα τμήματα τα οποία ενώνουν αυτά τα σημεία με τη σειρά που δίνονται. Μπορεί κανείς επίσης να καθορίσει το χρώμα, το πάχος και τη φωτεινότητα της γραμμής. Το χρώμα πρέπει να είναι σε δεκαεξαδική μορφή όπως στην HTML. Περισσότερες λεπτομέρειες για τα polylines μπορούν να αναζητηθούν στο documentation του Google Maps API.
- **Controls:** Για να χρησιμοποιήσουμε έλεγχο πάνω στο χάρτη, όπως μετακίνηση ή zoom ή οποιονδήποτε άλλο έλεγχο, υπάρχει η μέθοδος addControl. Το Maps API έχει ενσωματωμένους τους εξής ελέγχους που μπορεί κανείς να χρησιμοποιήσει στο χάρτη του:
 - **GLargeMapControl:** Ένας μεγάλου εύρους κίνησης/zoom έλεγχος που χρησιμοποιείται στο Google Maps.
 - **GSmallMapControl:** Ένας μικρότερος έλεγχος κίνησης/zoom που χρησιμοποιείται στο Google Local.
 - **GSmallZoomControl:** Ένας μικρός έλεγχος zoom που χρησιμοποιείται στο Google Maps για να εμφανίσει οδηγίες καθοδήγησης.
 - **GMapTypeControl:** Έλεγχος για να μπορεί ο χρήστης να αλλάζει τους διάφορους τύπους χαρτών (για παράδειγμα Map και Satellite).

- XML and RPC: Το Google Maps API προσφέρει μια «factory μέθοδο» για τη δημιουργία XmlHttpRequest αντικειμένων τα οποία «δουλεύουν» σε πρόσφατες εκδόσεις των Internet Explorer, Firefox, και Safari. Δίνεται επίσης η δυνατότητα να κάνουμε parsing σε ένα XML έγγραφο με τη στατική μέθοδο GXml.parse, η οποία παίρνει σαν είσοδο ένα XML string. Να σημειώσουμε μόνο ότι το Google Maps API δεν απαιτεί τη χρήση XML ή XmlHttpRequest για να λειτουργήσει καθώς βασίζεται αποκλειστικά σε ένα «καθαρό» JavaScript/DHTML API.

5.3.2 To “Hello World” στο Google Maps

Ο πιο εύκολος τρόπος για να κατανοήσει κάποιος το Google Maps API είναι με το να μελετήσει ένα απλό παράδειγμα 88. Η παρακάτω ιστοσελίδα απεικονίζει έναν χάρτη 500x300 με κέντρο το Πάλο Άλτο στην Καλιφόρνια (βλ. παρ 5.3.1).

```
<!DOCTYPE html "-//W3C//DTD XHTML 1.0 Strict//EN"
    "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
  <head>
    <meta http-equiv="content-type"
content="text/html; charset=utf-8"/>
    <title>Google Maps JavaScript API
Example</title>
    <script src="http://maps.google.com/maps?
file=api&v=2&key=abcdefghijklmnop"
        type="text/javascript"></script>
    <script type="text/javascript">

function initialize() {
  if (GBrowserIsCompatible()) {
    var map = new
GMap2(document.getElementById("map_canvas"));
    map.setCenter(new GLatLng(37.4419,
-122.1419), 13);
    map.openInfoWindow(map.getCenter(),
document.createTextNode("Hel-
lo, world"));
  }
}

</script>
</head>
<body onload="initialize()"
onunload="GUnload()">
  <div id="map_canvas" style="width: 500px;
height: 300px"></div>
</body>
```

Στο απλό αυτό παράδειγμα, υπάρχουν πέντε βασικά βήματα :

1. Τοποθετούμε τον κώδικα γραμμένο στην γλώσσα JavaScript του Google Maps API μέσα στην επικεφαλίδα *script*.
2. Δημιουργούμε ένα στοιχείο *div* με το όνομα “map_canvas” το οποίο περιέχει τον χάρτη
3. Γράφουμε μια συνάρτηση στην γλώσσα JavaScript που να δημιουργεί

ένα αντικείμενο “map”

4. Δίνουμε ένα γεωγραφικό σημείο που να αποτελεί το κέντρο του χάρτη.

5. Αρχικοποιούμε το χάρτη με το γεγονός “*onLoad*” της επικεφαλίδας “*body*”.

Τα βήματα αυτά επεξηγούνται στην συνέχεια.

5.3.2.1 Loading to Google Maps API

```
<script src="http://maps.google.com/maps?file=api&v=2&key=abcdefg"
  type="text/javascript">
</script>
```

Το URL “http://maps.google.com/maps?file=api&v=2&key=abcdefg” δείχνει την τοποθεσία ενός αρχείου γραμμένο σε JavaScript, που περιέχει όλα τα σύμβολα και τους ορισμούς που κάποιος χρειάζεται, ώστε να χρησιμοποιήσει το Google Maps API. Κάθε τέτοια σελίδα πρέπει να περιέχει μία επικεφαλίδα που να δείχνει σε αυτό το URL, χρησιμοποιώντας το κλειδί που έχει πάρει από την εγγραφή του από το API. Στο παράδειγμα αυτό το κλειδί που χρησιμοποιείται είναι το “abcdefg”.

5.3.2.2 Τα στοιχεία DOM του χάρτη

```
<div id="map_canvas" style="width: 500px; height: 300px"></div>
```

Ο χάρτης που απεικονίζεται στην ιστοσελίδα πρέπει να τοποθετηθεί σε ένα σημείο της. Έτσι, δημιουργούμε ένα στοιχείο “*div*” και χρησιμοποιούμε την τεχνολογία του **DOM** (document object model).

Στο προηγούμενο παράδειγμα, δημιουργήσαμε ένα στοιχείο “*div*” με το όνομα “*map_canvas*” και καθορίζουμε το μέγεθος του, χρησιμοποιώντας τις ιδιότητες του στυλ του **DOM**. Αν δεν καθορίσουμε το μέγεθος του χάρτη από το **GMapOptions** από τον κατασκευαστή, ο χάρτης παίρνει το μέγεθος του τμήματος που το περιέχει.

5.3.2.3 GMap2 – Το κύριο αντικείμενο

```
var map = new GMap2(document.getElementById("map_canvas"));
```

Η **GMap2** είναι η κλάση που παριστάνει τον χάρτη. Ένα αντικείμενο από αυτή την κλάση δημιουργεί έναν χάρτη σε μια σελίδα.

Όταν δημιουργούμε έναν καινούργιο χάρτη, δημιουργούμε ένα κόμβο **DOM** σε μια ιστοσελίδα (συνήθως ένα *div* στοιχείο) για να περιέχει τον χάρτη. Όλοι οι κόμβοι **HTML** είναι παιδιά του αντικειμένου *document*, και μπορούμε να αναφερθούμε σε αυτούς με την μέθοδο *document.getElementById()*. Ο προηγούμενος κώδικας ορίζει μία μεταβλητή με το όνομα *map* και αναφέρεται σε ένα νέο αντικείμενο **GMap2**. Η συνάρτηση **GMap2()** ονομάζεται κατασκευαστής και ο ορισμός του φαίνεται παρακάτω.

Κατασκευαστής	Περιγραφή
GMap2(container, opt?)	Δημιουργεί έναν νέο χάρτη σε ένα συγκεκριμένο HTML στοιχείο, που είναι συνήθως ένα DIV στοιχείο. Μπορούμε επίσης να δώσουμε διάφορες παραμέτρους τύπου GMap2Options στην παράμετρο opts

Πίνακας 18 – Το αντικείμενο GMap2

5.3.2.4 Αρχικοποίηση του χάρτη

```
map.setCenter(new GLatLng(37.4419, -122.1419), 13);
```

Αφού έχουμε δημιουργήσει τον χάρτη, μένει μόνο ένα βήμα αυτό της αρχικοποίησης. Αυτή η αρχικοποίηση συνδέεται με την μέθοδο **setCenter()**. Η **setCenter()** χρειάζεται τις συντεταγμένες τύπου **GLatLng** και το επίπεδο τις εστίασης και αυτή η μέθοδος πρέπει να εκτελεστεί πριν εκτελεστεί οποιαδήποτε άλλη.

5.3.2.5 Loading the Map

```
<body onload="initialize()" onunload="GUnload()">
```

Θέλουμε ο χάρτης να τοποθετείται στην σελίδα αφού έχει φορτωθεί ολόκληρη η σελίδα, για αυτό εκτελούμε την συνάρτηση που χτίζει το αντικείμενο **GMap2** όταν το στοιχείο **<body>** της **HTML** σελίδες δέχεται ένα γεγονός “**onLoad**”. Κάτι τέτοιο εμποδίζει οποιαδήποτε ανεπιθύμητοι συμπεριφορά και μας δίνει την δυνατότητα για μεγαλύτερο έλεγχο για το πως και πότε σχεδιάζεται ο χάρτης.

Η συνάρτηση **GUnload()** είναι σχεδιασμένη για να αποφεύγονται οι περιπτώσεις όπου δεν υπάρχει ελεύθερη μνήμη.

5.3.2.6 Latitudes και Longitudes

Αφού έχουμε δημιουργήσει τον χάρτη χρειαζόμαστε ένα τρόπο ώστε να μπορούμε να αναφερθούμε σε διάφορες τοποθεσίες στο χάρτη. Το αντικείμενο **GLatLng** προσφέρει έναν τέτοιο μηχανισμό για το Google Maps API.

```
var myGeographicCoordinates = new GLatLng(myLatitude, myLongitude)
```

5.3.2.7 Παράθυρα Πληροφοριών – Info Window

Κάθε χάρτης του Google Maps API μπορούν να περιέχουν και ένα «**Παράθυρο Πληροφοριών**», το οποίο απεικονίζεται με ένα μικρό παράθυρο πάνω στον χάρτη. Το **Παράθυρο Πληροφοριών** μοιάζει με τα παράθυρα διαλόγων που υπάρχουν στα κόμιξ.

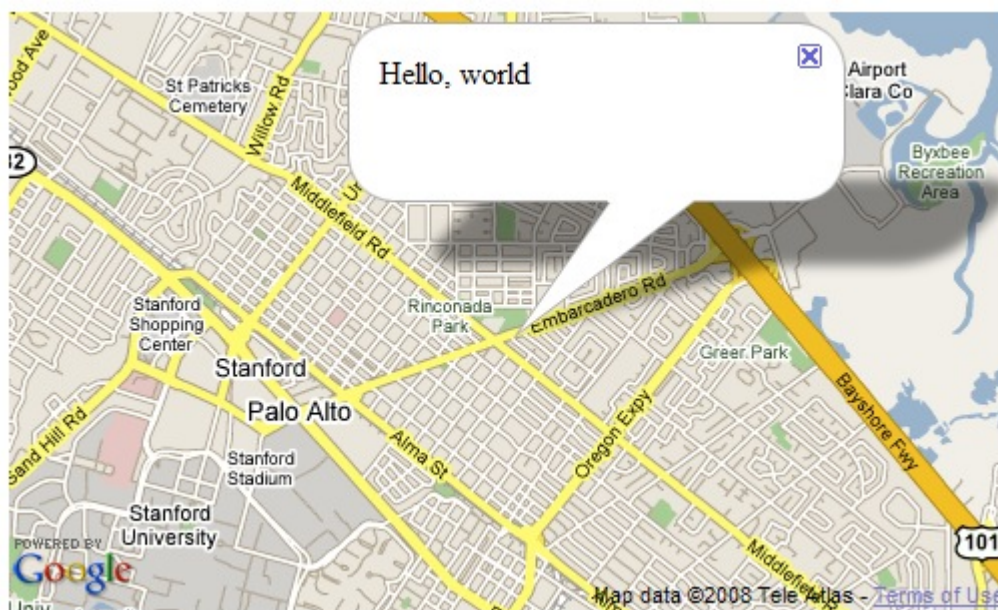
Το αντικείμενο **GMap2** provide την μέθοδο **openInfoWindow**, η οποία παίρνει σαν παραμέτρους ένα σημείο και ένα στοιχείο **HTML DOM**. Το στοιχείο **HTML DOM** είναι

συγγενής με το στοιχείο που περιέχει το *Παράθυρο Πληροφοριών*, και το *Παράθυρο Πληροφοριών* τοποθετείται στο συγκεκριμένο σημείο.

Ο ακόλουθος κώδικας απεικονίζει ένα *Παράθυρο Πληροφοριών* που είναι τοποθετημένο στο κέντρο του χάρτη και περιέχει την φράση *“Hello, world”*.

```
map.openInfoWindow(map.getCenter(),  
document.createTextNode("Hello, world"));
```

Στο επόμενο σχήμα απεικονίζεται ο χάρτης που κατασκευάσαμε



Σχήμα 5.19 - To Hello World του Google Map API

5.3.3 Οι Google maps της εφαρμογής μας

Στην παράγραφο 4.3.3 περιγράφηκαν οι λόγοι για τους οποίους χρησιμοποιήσαμε τους χάρτες Google. Εδάφιο αυτό θα αναφερθούμε κυρίως στις λεπτομέρειες υλοποίησης τους ώστε να συνεργάζονται άψογα με την εφαρμογή GeoJoy.

Επειδή υπάρχουν τρεις διαφορετικές περιπτώσεις χρησιμοποίησης των χαρτών, η κάθε μία με διαφορετικές παραμέτρους και λειτουργίες, κατασκευάσαμε τρεις διαφορετικές ιστοσελίδες που περιλαμβάνουν αυτούς του χάρτες. Για την δημιουργία των χαρτών χρησιμοποιήθηκε το αντικείμενο Gmap2 (Πίνακας 18 – Το αντικείμενο GMap2).

```
map.addControl(new GOverviewMapControl());  
map.addControl(new GLargeMapControl());
```

Επίσης, χρησιμοποιήθηκαν κάποια στοιχεία στην εμφάνιση των χαρτών οι “έλεγχοι” (controls), που επιτρέπουν την αλληλεπίδραση του χρήστη με τους χάρτες. Για την παρούσα εφαρμογή, χρησιμοποιήθηκαν τα GOverviewMapControl (αποτελεί ένα παράθυρο

προεσκόπισης που εμφανίζεται στην κάτω δεξιά γωνία του χάρτη) και `GLargeMapControl` (έλεγχος της εστίασης στον χάρτη. Εμφανίζεται στην πάνω αριστερή γωνία του χάρτη).

Στην περίπτωση των χαρτών που χρησιμοποιούνται για να παρουσιάσουν την γεωγραφική πληροφορία, ορίσαμε σαν κέντρο τους το σημείο στο οποίο αναφέρονται. Στο σημείο αυτό εμφανίζεται και ένας “δείκτης” (marker – στον χάρτη του Σχήμα 4.16 – Ένας χάρτης Google εμφανίζεται με κόκκινο χρώμα). Γενικά, οι “δείκτες” είναι σταθεροί, όπως στον χάρτη για την απλή απεικόνιση της πληροφορίας, όμως για την δεύτερη περίπτωση, όπου ο χρήστης μπορεί να αλλάξει τις συντεταγμένες του σημείου, προτιμήθηκε ο “δείκτης” να μπορεί να μετακινηθεί (draggable Marker) στο σημείο που θέλει ο χρήστης ώστε να διευκολύνεται αρκετά η διαδικασία της αλλαγής των συντεταγμένων.

```
var marker = new GMarker(point, {draggable: true});
GEvent.addListener(marker, "dragstart", function() {map.closeInfoWindow();});
GEvent.addListener(marker, "dragend", function() {
    marker.openInfoWindowHtml('<div>'+unescape(name)+'</div>'
+'<br>If this place press OK..');
    newPoint = marker.getLatLng();
    findCoord();
});
```

Κώδικας για τον Draggable Marker

Επίσης, σε αυτούς τους χάρτες εμφανίζεται και ένα “παράθυρο πληροφοριών” (info window –Σχήμα 4.7), το οποίο περιέχει το όνομα της γεωγραφικής πληροφορίας στην οποία αντιστοιχεί ο χάρτης αυτός.

Στην ενέργεια που ο χρήστης προσθέτει κάποια γεωγραφική πληροφορία που δεν βρέθηκε από την διαδικασία απεικόνισης, πρέπει η εισαγωγή των συντεταγμένων της πληροφορίας να γίνεται με πολύ απλό τρόπο. Έτσι όταν γίνει η αίτηση για προσθήκη, ανοίγει ένα νέο παράθυρο με ένα χάρτη που απεικονίζει την Ελλάδα. Σε αυτό ο χρήστης εντοπίζει το σημείο στον χάρτη και ενεργοποιεί του δεξιά κλικ menu του χάρτη (right-click menu) και εκτελεί την ενέργεια “geotag this point”. Τα περιεχόμενα και η εμφάνιση αυτού του menu, γίνονται με την βοήθεια της τεχνολογίας του DOM και η πυροδότηση του γίνεται με την ενεργοποίηση του δεξιού κλικ γεγονόςτος (right-click event). Το αντικείμενο `GEvents` είναι αυτό που χειρίζεται τα γεγονότα στο Google maps API.

```
// Δημιουργία του κόμβου που θα περιέχει το κείμενο “Geotag this point” με
// το style που φαίνεται παρακάτω και το ορίζουμε σαν παιδί του “map”
contextmenu = document.createElement("div");
contextmenu.style.visibility="hidden";
contextmenu.style.background="#ffffff";
contextmenu.style.border="1px solid #8888FF";
a = document.createElement("a");
```



```

a.setAttribute("href", "javascript:findCoord()");
a.textContent="Geotag this point" ;
contextmenu.appendChild(a);
map.getContainer().appendChild(contextmenu);
// Ορισμός του Event listener για το right-click event
// Η συνάρτηση που καλείται εμφανίζει το κείμενο “geotag this point” δίπλα // στο σημείο που έγινε
η πυροδότηση του γεγονότος και εντοπίζει τις
//συντεταγμένες του σημείου.
GEvent.addListener(map,"singlerightclick",function(pixel,src,overlay){
    clickedPixel = pixel;
    var x=pixel.x;
    var y=pixel.y;
    if (x > map.getSize().width - 120){
        x = map.getSize().width - 120
    }
    if (y > map.getSize().height - 100) {
        y = map.getSize().height - 100
    }
    var pos = new GControlPosition(G_ANCHOR_TOP_LEFT, new GSize(x,y));
    lat_lgn = map.fromDivPixelToLatLng(pixel);
    pos.apply(contextmenu);
    contextmenu.style.visibility = "visible";
});

```

Κώδικας για την Προσθήκη Γεωγραφικής

Πληροφορίας

5.4 JavaScript

Η JavaScript 888είναι γλώσσα προγραμματισμού η οποία έχει σαν σκοπό την παραγωγή δυναμικού περιεχομένου σε ιστοσελίδες. Η JavaScript είναι client side. Κατά την κατασκευή της επηρεάστηκε από πολλές γλώσσες ενώ σχεδιάστηκε ώστε να μοιάζει με την Java, αλλά είναι πιο εύκολο για κάποιον που δεν είναι προγραμματιστής να δουλέψει με αυτήν. Έγινε γνωστή λόγω της συχνής της χρήσης σε ιστοσελίδες (client-side JavaScript) αλλά χρησιμοποιείται και σε περιπτώσεις που απαιτείται ενσωματωμένος κώδικας (π.χ. Microsoft Gadgets στο Sidebar των Windows Vista). Παρόλο το όνομα της, δεν έχει καμία απολύτως σχέση με την Java, αν και οι δύο έχουν παρόμοια σύνταξη και η JavaScript έχει τις ίδιες ονομασίες στα αντικείμενα και τους ίδιους κανόνες στις ονομασίες των μεταβλητών.

Client-side σημαίνει ότι η επεξεργασία του κώδικα JavaScript και η παραγωγή του τελικού περιεχομένου HTML δεν πραγματοποιείται στον server, αλλά στο πρόγραμμα περιήγησης των επισκεπτών. Η JavaScript δεν έχει καμία απαίτηση από πλευράς δυνατοτήτων του server για να εκτελεστεί (επεξεργαστική ισχύ, συμβατό λογισμικό διακομιστή), αλλά βασίζεται στις δυνατότητες του browser των επισκεπτών. Επίσης μπορεί να ενσωματωθεί σε στατικές σελίδες HTML. Για τους παραπάνω λόγους, είναι και η βασική γλώσσα για την δημιουργία extensions.

5.5 Postgres SQL

Η PostgreSQL 8.8 είναι ένα αντικειμενοστραφές σύστημα διαχείρισης Βάσεων Δεδομένων (ORDBMS). Εδώ χρησιμοποιήθηκε η PostgreSQL 8.2 καθώς και το pgAdmin (PostgreSQL Tools) v. 1.6.3 για τη δημιουργία της κεντρικής βάσης δεδομένων.

Για την υλοποίηση μας χρειάστηκε μόνο ένας πίνακας για την δημιουργία της βάσης. Ο πίνακας αυτός εκτός από το όνομα της γεωγραφικής πληροφορίας περιέχει στοιχεία που περιγράφουν την θέση της στο κείμενο της ιστοσελίδας δηλ. την ενθυλάκωση της φράσης, στην οποία ανήκει, μέσα στο HTML κείμενο και την θέση της μέσα στην φράση αυτή. Επίσης περιέχει την διεύθυνση της ιστοσελίδας στην οποία βρέθηκε αυτή η γεωγραφική πληροφορία, ένα id για να προσφέρει μοναδικότητα σε κάθε εγγραφή και μια μεταβλητή που μας πληροφορεί αν η εγγραφή έγινε με αυτόματο ή χειρονακτικό τρόπο ή αν άλλαξαν τα πεδία της ή αν διαγράφηκε. Τέλος, περιέχει τρία πεδία που βοηθούν στην απεικόνιση της σε ένα Google map, το γεωγραφικό πλάτος και ύψος και το επίπεδο της εστίασης στον χάρτη.

GeoJoy Table					
geo_name	geo_url	geo_id	geo_path	geo_start_point	geo_end_point
Character varying [70]	Character varying [300]	Integer	Character varying [300]	integer	integer

GeoJoy Table (συνέχεια)			
geo_lat	geo_lng	geo_zoom	geo_type_added
Double precision	Double precision	integer	integer

5.6 Tomcat

Ο Tomcat 8.8 είναι ένα servlet container που υποστηρίζει τις τεχνολογίες Java Servlets και JSPs καθώς και στατικές σελίδες δεχόμενος requests στην πόρτα 8080. Ο Apache Tomcat αναπτύσσεται σε ένα ανοικτό και συμμετέχον περιβάλλον και διανέμεται με άδεια

λογισμικού Apache. Ο Apache Tomcat προορίζεται να είναι μια συνεργασία από τους καλύτερους προγραμματιστές του κόσμου. Υποστηρίζει έναν μεγάλο αριθμό μεγάλης κλίμακας και κρίσιμων εφαρμογών του Ιστού. Στην παρούσα εφαρμογή χρησιμοποιήθηκε η **έκδοση 6.0**.

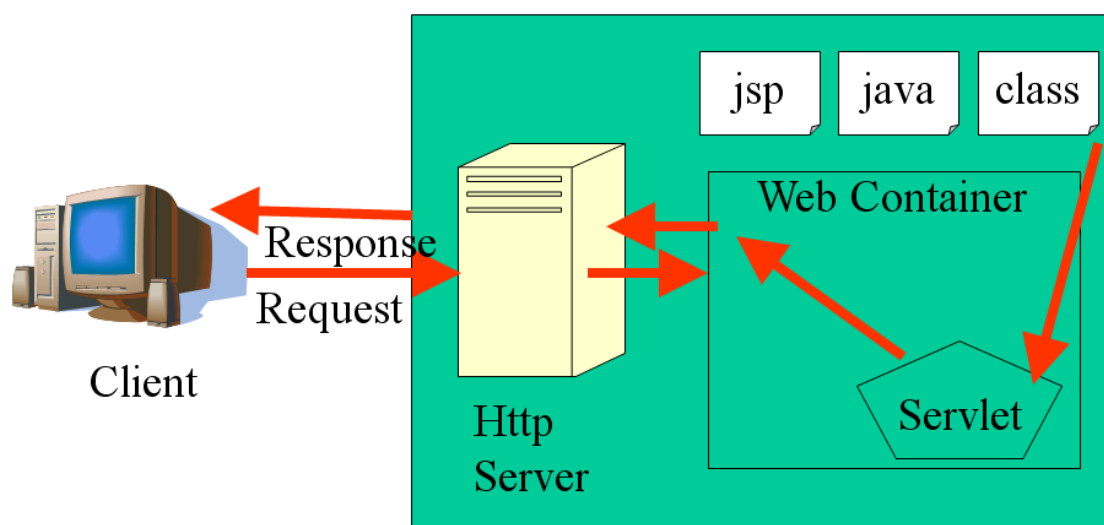
Ένα servlet μέσα στο Tomcat λειτουργεί ως εξής : Το **container** που διαθέτει το **Java Servlet** δημιουργεί ένα **στιγμιότυπο (instance)** την **πρώτη** φορά που γίνεται μία αίτηση γι' αυτό. Ουσιαστικά η κάθε αίτηση εξυπηρετείται από το **container** που χρησιμοποιεί ένα **νέο νήμα εκτέλεσης (thread)** για την εξυπηρέτηση της αίτησης από το **ίδιο** αντικείμενο που δημιουργήθηκε όταν έγινε η πρώτη αίτηση.

5.7 Servlets

Τα Servlets είναι προγράμματα της Java για την επέκταση και την ενίσχυση Web servers. Χρησιμοποιούνται σε εφαρμογές που απαιτούν **προεπεξεργασία**, πριν την αποστολή της απόκρισης στον αιτητή. Παρέχουν μια βασισμένη στα συστατικά, ανεξάρτητη πλατφόρμας μέθοδο για τις WEB εφαρμογές, χωρίς τους περιορισμούς απόδοσης των προγραμμάτων της CGI, είναι δηλαδή κατά κάποιο τρόπο CGIs αλλά γραμμένα σε **Java**, οπότε κληρονομούν όλα τα πλεονεκτήματα που προσφέρει αυτή η γλώσσα. Τα Servlets αποτελούν το **Server Side** πρόσωπο της Java σε αντίθεση με τα Applets που αποτελούν το **Client Side** πρόσωπο της Java.

Ένα servlet είναι μια κλάση όπως όλες οι άλλες με την διαφορά ότι κάνει implement το javax.servlet.Servlet interface.

Πρακτικά, τα περισσότερα servlets κάνουν extend την javax.servlet.http.HttpServlet, ενώ σπανιότερα κάνουν extend την javax.servlet.GenericServlet.



Σχήμα 5.8 – Η λειτουργία ενός Servlet

Τα βασικά χαρακτηριστικά των Servlets είναι :

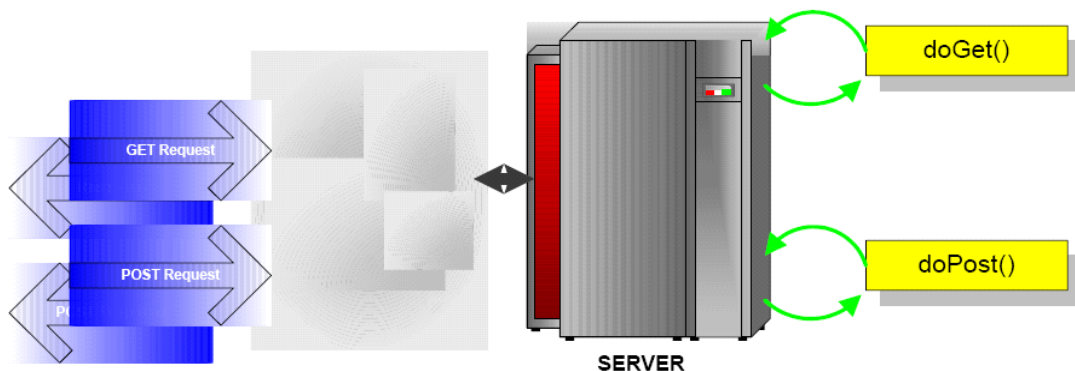
- **Ανεξάρτητα** από την πλατφόρμα και τον server - “Write once Serve Everywhere” - Portability
- **Εκμεταλλεύονται πλήρως** το Java API, rmi, corba, database connectivity.
- **Εύκολη προσθήκη** 3rd party components.
- **Αποδοτικότητα & Αντοχή** - Μένουν στην μνήμη μεταξύ διαδοχικών καλεσμάτων
- **Ασφάλεια Κώδικα** (Garbage Collector, Exception Handling)
- **Κομψότητα** (Elegance) (Object Oriented, Clean Code, Modular, Simple)
- **Εύκολο integration** νέων υπηρεσιών

Αλληλεπιδρούν με τους clients με ένα μοντέλο αίτησης-απάντησης. Ένα Java Servlet θεωρείται μία επέκταση των δυνατοτήτων εξυπηρετητών διαδικτύου (web servers), δηλαδή εξυπηρετητών που βασίζονται στο πρωτόκολλο HTTP και για να επικοινωνήσει με τον εξυπηρετητή χρειάζεται έναν container, στην περίπτωση μας τον Apache Tomcat. Ο **container server** (Tomcat) δέχεται την **αίτηση** από τον πελάτη, και **καλεί** το αντίστοιχο servlet (κλάση) **περνώντας** του και τις **παραμέτρους** του χρήστη. Τα Servlets η τεχνολογία της Java για τη δημιουργία δυναμικών σελίδων στο διαδίκτυο. Επίσης έχουν τη δυνατότητα να εμπεριέχουν HTML tags στον Java κώδικα. Σήμερα τα servlets είναι μια δημοφιλής επιλογή για την οικοδόμηση των διαλογικών εφαρμογών Ιστού.

Όλα τα servlets υλοποιούν το interface τους μέσω μιας εκ των δύο πιο κάτω βασικών κλάσεων:

- **GenericServlet** (javax.servlet package) - γενικά servlets ανεξάρτητα πρωτοκόλλου
- **HttpServlet** (javax.servlet.http package) - http servlets (χρήση αυτού στην εφαρμογή μας)

Όταν ένα servlet δέχεται ένα κάλεσμα από τον πελάτη (client), λαμβάνει δύο αντικείμενα (objects), ένα **ServletRequest**, που εξασφαλίζει την επικοινωνία από τον πελάτη προς τον server και ένα **ServletResponse**, που εξασφαλίζει την επικοινωνία από το servlet πίσω στον πελάτη, τα οποία είναι interfaces ορισμένα στο javax.servlet package. Οι βασικές μέθοδοι είναι α) η **doGet** και β) η **doPost**. Τέλος, ο κύκλος ζωής του είναι ίδιος σε κάθε servlet. Γεννιέται με την φόρτωση και την αρχικοποίηση του από τον server και κατά την διάρκεια της ζωής του δέχεται μηδέν ή και περισσότερα client **requests**. Ο θάνατος του πραγματοποιείται όταν ο server το κάνει **remove** (ορισμένοι servers εκτελούν αυτό το βήμα μόνο όταν κάνουν shut down).



Σχήμα 5.9 - HttpServlet

5.8 Java

Η Java 8, 88 είναι μία αντικειμενοστραφής γλώσσα προγραμματισμού που σχεδιάστηκε από την εταιρεία πληροφορικής *Sun Microsystems*. Ένα από τα βασικά πλεονεκτήματα της Java έναντι των περισσότερων άλλων γλωσσών είναι η ανεξαρτησία του λειτουργικού συστήματος και πλατφόρμας.

Στη παρούσα διπλωματική χρησιμοποιήθηκε η γλώσσα Java για την τροποποίηση του εργαλείου της αυτόματης γεωκωδικοποίησης και για την δημιουργία των Java Servlets. Χρησιμοποιήθηκε το περιβάλλον ανάπτυξης Eclipse (v. 3.3.1.1) και το JRE *System Library (jre 1.6.0_03)*. Για την περίπτωση των Java Servlets προστέθηκε το πακέτο *javax.servlet-api.jar v. 2.5*.

5.9 AJAX

Αντίθετα με τον παραδοσιακό πρότυπο της εφαρμογής Ιστού, όπου ο ίδιος ο browser είναι αρμόδιος για την έναρξη των αιτημάτων προς , και την επεξεργασία των αιτημάτων από, τον κεντρικό υπολογιστή δικτύου (Web Server), το πρότυπο **Ajax** παρέχει ένα ενδιάμεσο στρώμα - καλείται και ως μηχανή Ajax - για να χειριστεί αυτήν την επικοινωνία. Μια μηχανή Ajax είναι πραγματικά ακριβώς ένα αντικείμενο JavaScript ή μια λειτουργία που καλείται κάθε φορά που ζητούνται πληροφορίες από τον κεντρικό υπολογιστή. Σε αντίθεση με το παραδοσιακό μοντέλο της πραγματοποίησης μιας σύνδεσης με έναν άλλο πόρο (όπως μια άλλη ιστοσελίδα), στη μηχανή Ajax κάθε σύνδεση εκτελεί μια κλήση, η οποία σχεδιάζει και εκτελεί το αίτημα. Το αίτημα γίνεται ασύγχρονα, δηλαδή η εκτέλεση του κώδικα συνεχίζεται χωρίς να περιμένει απάντηση.

Ο κεντρικός υπολογιστής - που παραδοσιακά εξυπηρετεί HTML, εικόνες, css, ή JavaScript - διαμορφώνεται ώστε να επιστρέφει στοιχεία που η μηχανή Ajax μπορεί να χρησιμοποιήσει.

Αυτό το στοιχείο μπορεί να είναι σαφές κείμενο, XML, ή οποιοδήποτε άλλο σχήμα στοιχείων που μπορεί κάποιος να χρειαστεί. Η μόνη απαίτηση είναι η μηχανή Ajax μπορεί να καταλάβει και να ερμηνεύσει αυτά τα στοιχεία.

Όταν η μηχανή Ajax λάβει την απάντηση από τον κεντρικό υπολογιστή(server), ξεκινάει την επεξεργασία των πληροφοριών που της παρήχθησαν και εκτελώντας συνήθως αλλαγές στο user interface. Επειδή αυτή η διαδικασία περιλαμβάνει τη μεταφορά λιγότερων πληροφοριών από το παραδοσιακό πρότυπο εφαρμογής Ιστού, οι αλλαγές στο user interface είναι γρηγορότερες, και ο χρήστης είναι σε θέση να κάνει την εργασία του/της γρηγορότερα. Το σχήμα 1-1 επιδεικνύει τη διαφορά μεταξύ των παραδοσιακού και Ιστού Ajax προτύπων εφαρμογής.

5.10 PHP

Η **PHP** 88, της οποίας τα αρχικά αντιπροσωπεύουν το "*PHP: Hypertext Preprocessor*" είναι μια ευρέως χρησιμοποιούμενη, ανοιχτού κώδικα, γενικού σκοπού scripting γλώσσα προγραμματισμού, η οποία είναι ειδικά κατάλληλη για ανάπτυξη εφαρμογών για το Web και μπορεί να ενσωματωθεί στην HTML. Η σύνταξη της παίρνει στοιχεία των C, Java, και Perl και είναι εύκολη στην μάθηση. Ο κύριος στόχος της γλώσσας είναι να επιτρέπει σε web developers να γράφουν δυναμικά παραγόμενες σελίδες (Web Pages) γρήγορα, αλλά κανείς μπορεί να κάνει πολύ περισσότερα με την **PHP**.

Η **PHP** στην παρούσα διπλωματική αποτελεί την γέφυρα για την μεταφορά δεδομένων από το extension στην κεντρική βάση δεδομένων. Έτσι χρησιμοποιήθηκε ο **HTTP Server Apache v.2.0.55 και η PHP v.5.1.1**.

6

Παρουσίαση

Αποτελεσμάτων

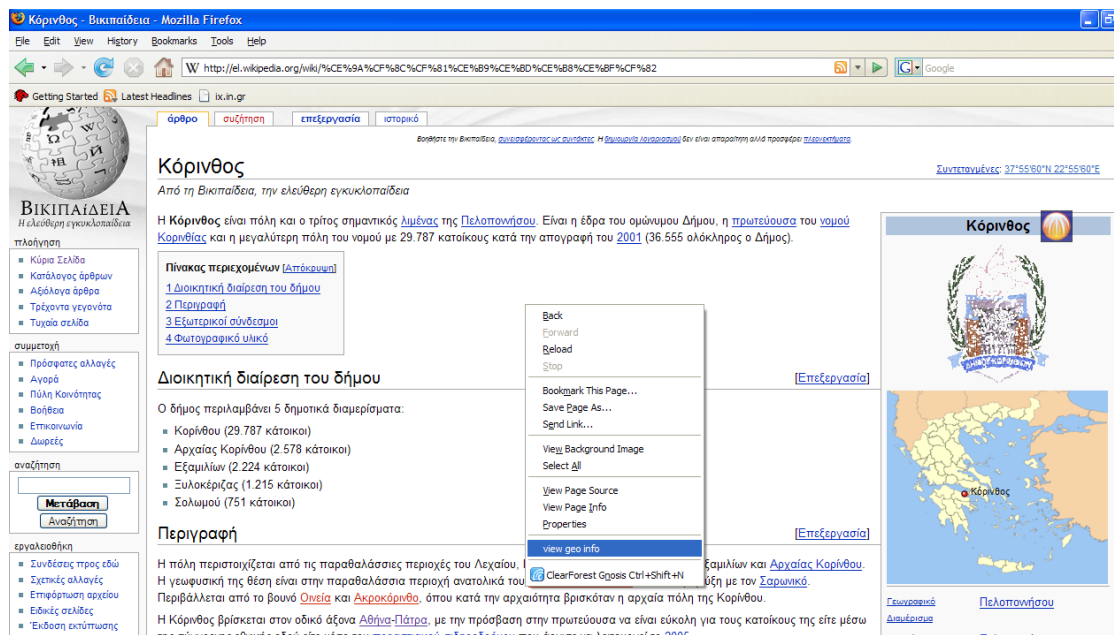
Το σύστημα αυτό λόγω της ιδιαιτερότητας του να δέχεται δεδομένα από τον χρήστη, είναι εύλογο να απευθύνεται σε ιστοσελίδες που δεν τροποποιούνται πολύ συχνά ή σχεδόν ποτέ, όπως λ.χ. τις Βικιπαιδείας. Όμως, τα παραπάνω δεν αποτελούν λόγο να μην χρησιμοποιηθεί σε οποιαδήποτε ιστοσελίδα. Ακολουθούν κάποια παραδείγματα χρήσης του και ένας οδηγός χρήσης.

6.1 Οδηγός Χρήσης

Στην ενότητα αυτή περιγράφεται ο τρόπος με τον οποίο ο κάθε χρήστης θα μπορεί να χρησιμοποιήσει το σύστημα. Η εγκατάσταση του είναι αρκετά απλή, αρκεί να εγκαταστήσει το αντίστοιχο extension στο φυλομετρητή Mozilla FireFox.

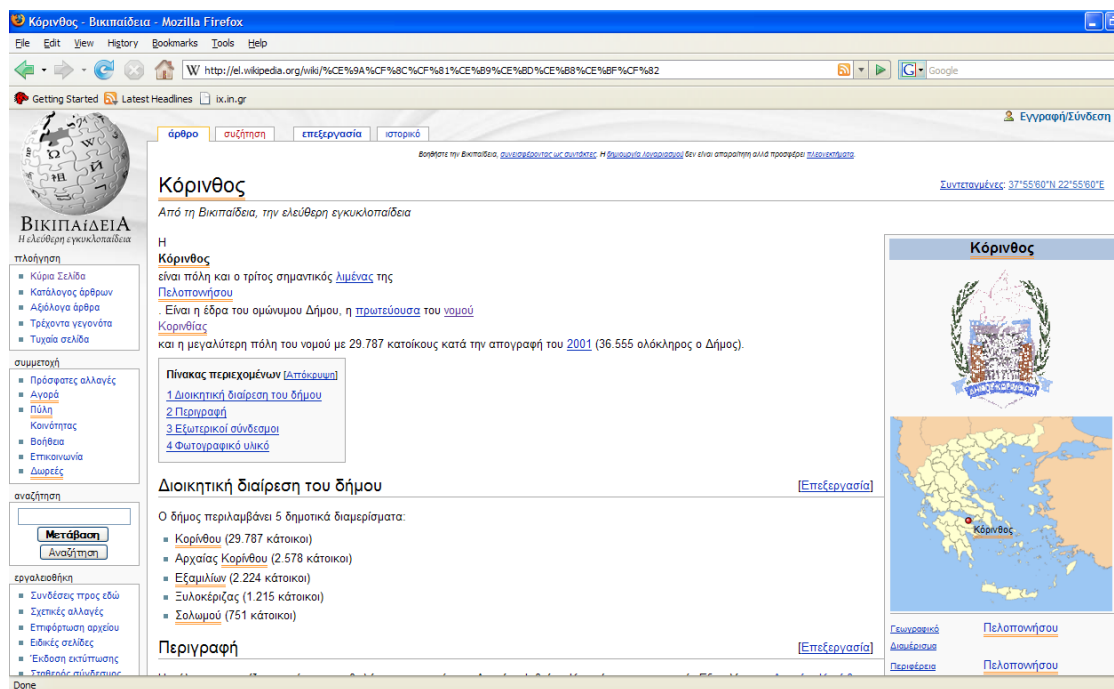
6.1.1 Απεικόνιση Γεωγραφικής πληροφορίας

Όταν ο χρήστης ανακτήσει στον Browser την ιστοσελίδα που επιθυμεί, εκτελεί την εντολή “view geo info” από το *right-click menu* του browser (Σχήμα 6.20).



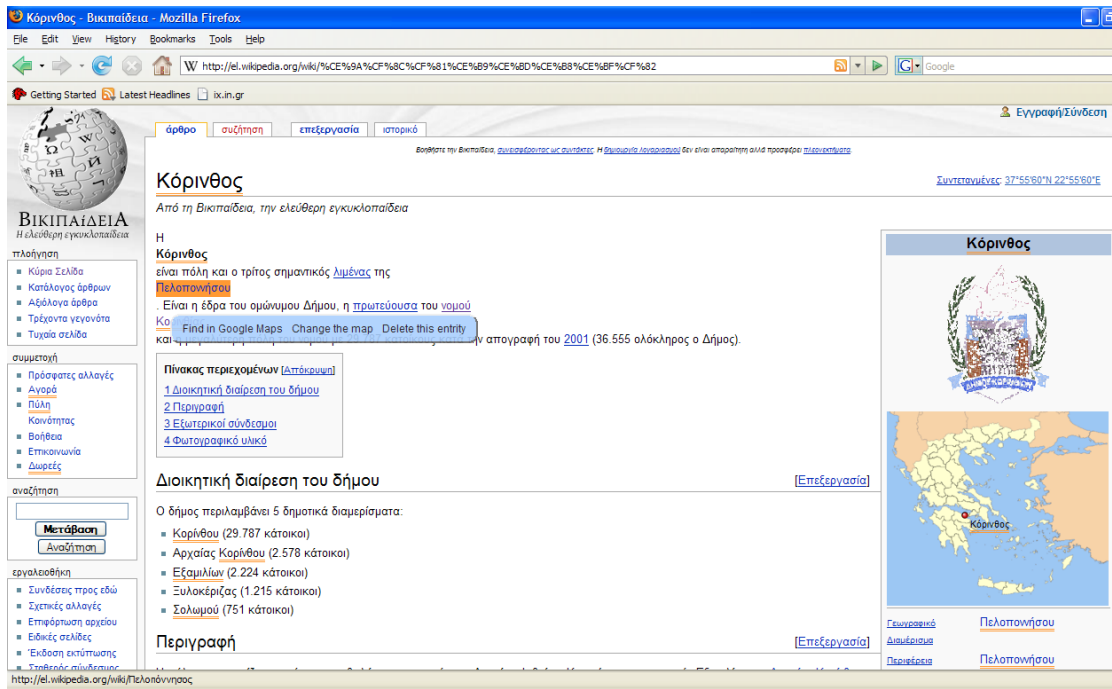
Σχήμα 6.20

Μέσα σε λίγα δευτερόλεπτα όλες οι γεωγραφικές πληροφορίες της ιστοσελίδας έχουν απεικονιστεί (Σχήμα 6.21).



Σχήμα 6.21

Τοποθετώντας απλά το ποντίκι πάνω σε μια από τις υπογραμμισμένες οντότητες εμφανίζεται το *hover – menu* που της αντιστοιχεί (Σχήμα 6.22) και για να δει που βρίσκεται στον χάρτη απλά εκτελεί την εντολή “*view in Google maps*” και εμφανίζεται ο σχετικός χάρτης.



Σχήμα 6.22

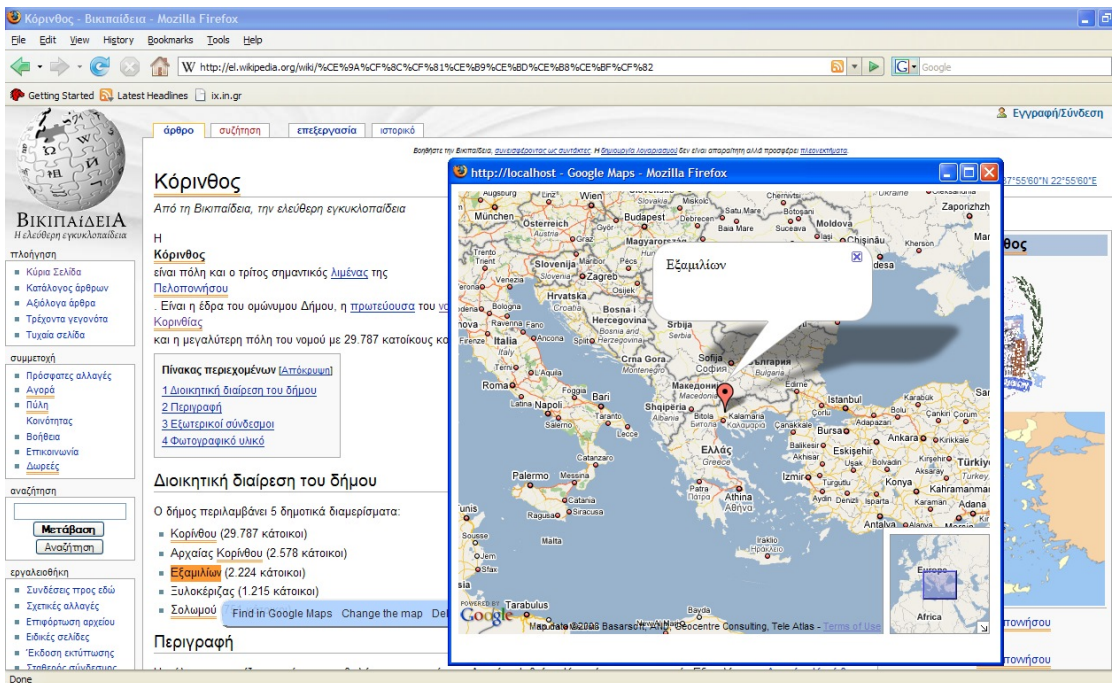
6.1.2 Διόρθωση Λαθών

Η προηγούμενη διαδικασία έχει βοηθήσει στην απεικόνιση των γεωγραφικών οντοτήτων της ιστοσελίδας. Αν ο χρήστης εντοπίσει λάθη ή παραλήψεις στα δεδομένα μπορεί να τα διορθώσει ή και να προσθέσει καινούργια.

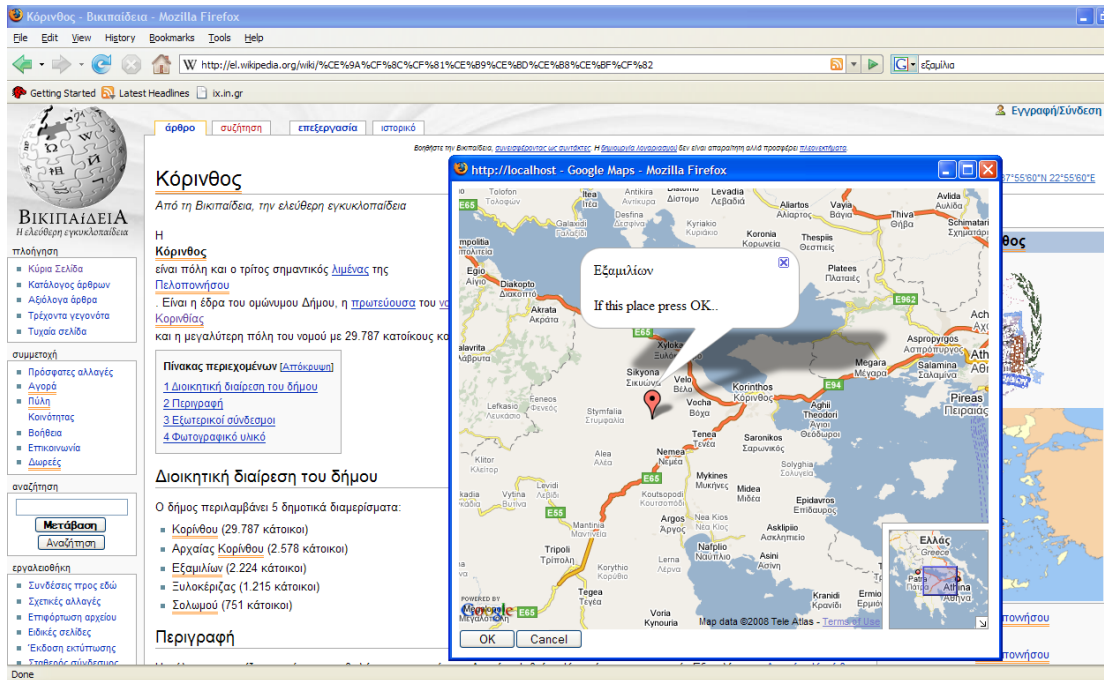
6.1.2.1 Αλλαγή Συντεταγμένων

Από το hover-menu που αντιστοιχεί σε κάθε οντότητα, ο χρήστης εκτελεί την εντολή “Change The Map” και ο σχετικός χάρτης εμφανίζεται (Σχήμα 6.23).

Σχήμα 6.23



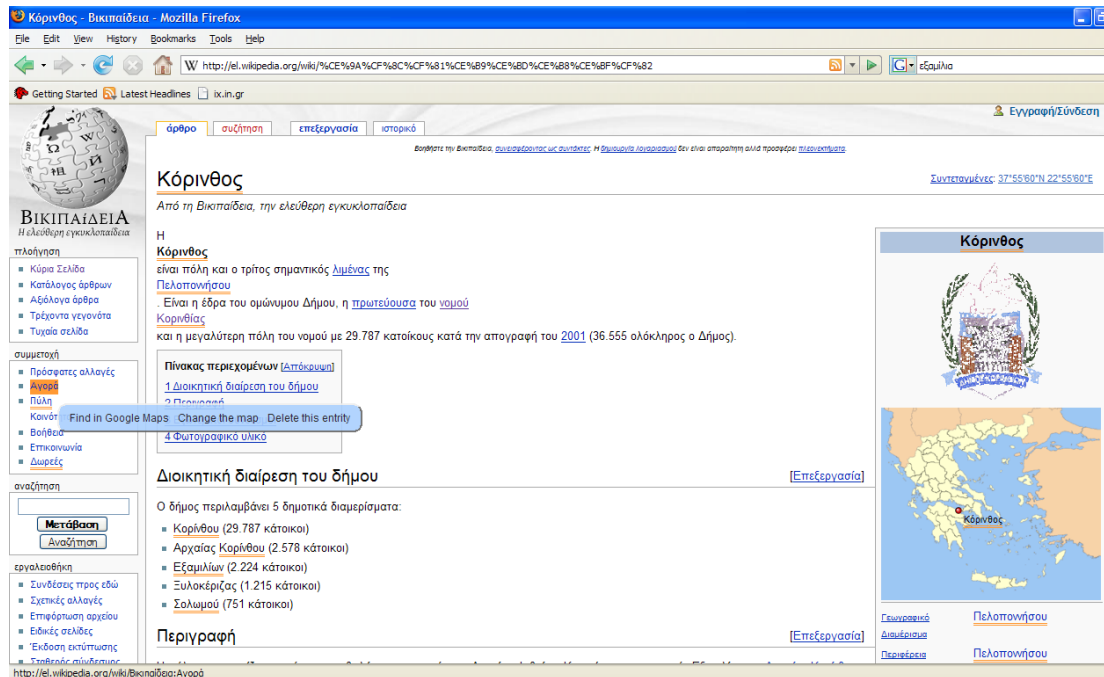
Το σημείο που βρίσκεται η οντότητα στον χάρτη είναι σημειωμένο με ένα δείκτη Αν ο χρήστης θεωρήσει πως ο δείκτης είναι τοποθετημένος λανθασμένα, απλά θα μετακινήσει το δείκτη στο σημείο που επιθυμεί και θα πατήσει το “OK” που θα εμφανιστεί στο κάτω δεξί μέρος του χάρτη, ώστε να αποθηκευτούν οι αλλαγές στην κεντρική βάση δεδομένων (Σχήμα 6.24).



Σχήμα 6.24

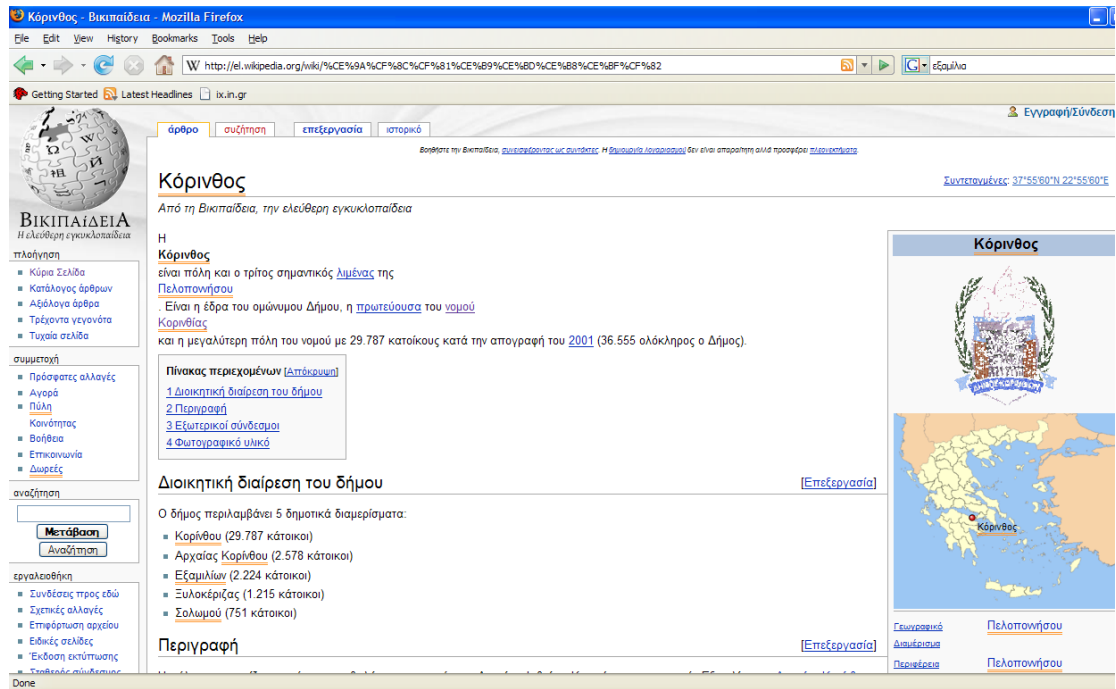
6.1.2.2 Διαγραφή Γεωγραφικής Οντότητας

Αν στην ιστοσελίδα έχουν υπογραμμιστεί λέξεις ή εκφράσεις που δεν αποτελούν κάποια γεωγραφική πληροφορία, ο χρήστης από το *hover-menu* τους μπορεί να εκτελέσει την εντολή “delete” ώστε να τις διαγράψει από την κεντρική βάση δεδομένων (Σχήμα 6.25). Αφαιρείται η οντότητα “Αγορά”.



Σχήμα 6.25

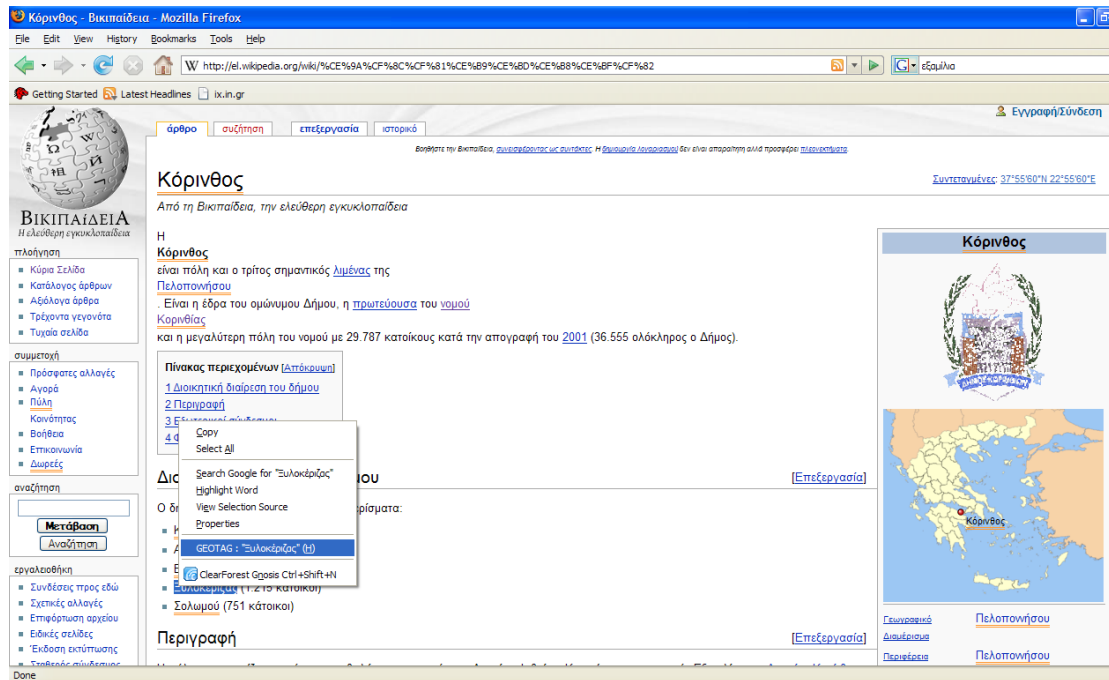
Στο Σχήμα 6.26 φαίνεται η ιστοσελίδα τροποποιημένη. Η οντότητα “Αγορά” δεν είναι πλέον σημειωμένη.



Σχήμα 6.26

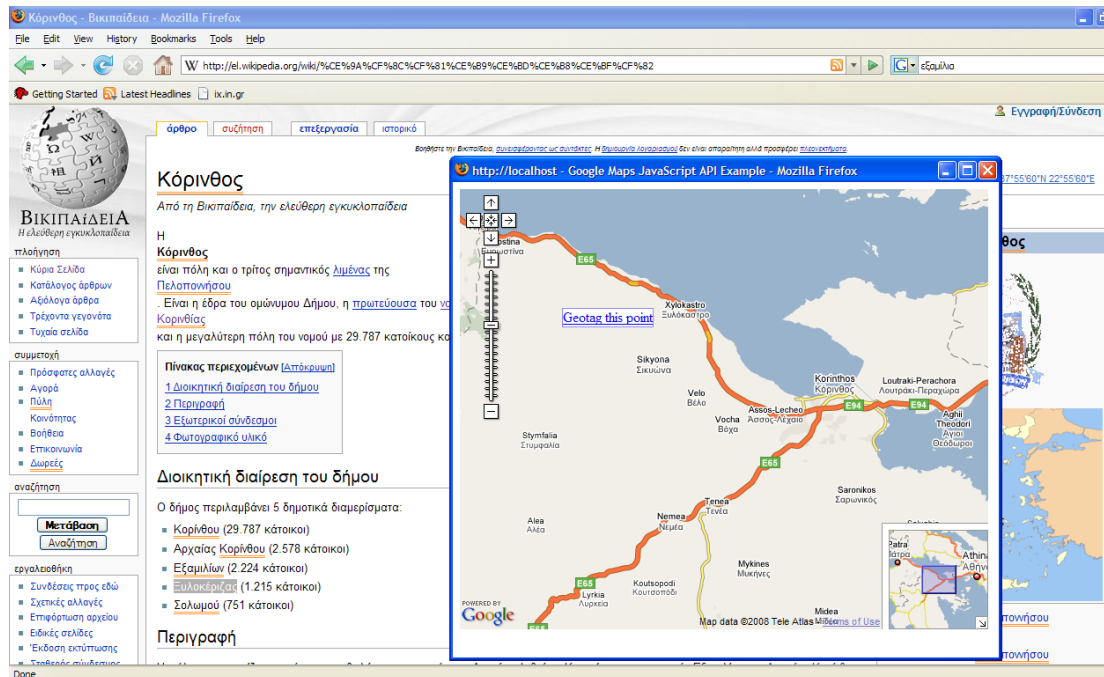
6.1.2.3 Προσθήκη Νέας Γεωγραφικής Οντότητας

Σε περίπτωση που υπάρξει κάποια γεωγραφική πληροφορία που δεν έχει απεικονιστεί, ο χρήστης μαρκάρει την λέξη ή την έκφραση αυτή και εκτελεί την εντολή “geotag : ...” (Σχήμα 6.27) από το *right-click menu* του browser.



Σχήμα 6.27


Στο χάρτη που εμφανίζεται, ο χρήστης βρίσκει το σημείο στον χάρτη που πιστεύει ότι ανήκει η πληροφορία αυτή και εκτελεί την εντολή “geotag” από το *right-click menu* του χάρτη και μετά “OK” για να αποθηκευτούν οι αλλαγές στην κεντρική βάση δεδομένων (Σχήμα 6.28).



Σχήμα 6.28

6.2 Μερικά Παραδείγματα

Χρησιμοποιούμε το εργαλείο που κατασκευάσαμε και διορθώνουμε ότι λάθη υποπέσουν στην αντίληψή μας στις δύο ιστοσελίδες που χρησιμοποιήθηκαν στα παραδείγματα της ενότητας 3.3.



ΒΙΚΙΠΑΙΔΕΙΑ
Η ελεύθερη εγκυκλοπαίδεια

Εγγραφή/Σύνδεση

άρθρο αναζήτηση επεξεργασία ιστορικό

βοήθεια της Βικιπαίδειας · [συμμετοχές/εργασίες](#) · [η Βικιπαίδεια για αρχάριους](#) · [έναν στην απιστία αλλά προσηύχοντας](#)

Κόρινθος

Από τη Βικιπαίδεια, την ελεύθερη εγκυκλοπαίδεια

Συντεταγμένες: 37°55′50″N 22°55′50″E

πλοήγηση

- Κύρια Σελίδα
- Κατάλογος άρθρων
- Αξιολόγητα άρθρα
- Τρέχοντα γεγονότα
- Τυχαιά σελίδα

συμμετοχή

- Πρόσφατες αλλαγές
- Αγορά
- Πύλη Κοινότητας
- Βοήθεια
- Επικοινωνία
- Διωρείς

αναζήτηση

εργαλειοθήκη

- Συνδέσεις προς εδώ
- Σχετικές αλλαγές
- Εμφάνιση αρχείου
- Εθνικές σελίδες
- Εξόδοση επιπτώσεων
- Σχεδιασμός συνδέσεων
- Βιβλιογραφικές παραπομπές

άλλες γλώσσες

- العربية
- Català
- Česky
- Словенски
- Dansk
- Deutsch
- English
- Esperanto
- Español
- Suomi
- Français
- געווי
- Magyar
- Italiano
- 日本語
- བོད་སྐད་
- Latina
- Lumbaart
- Lietuvių
- Nederlands
- Norsk (nynorsk)
- Norsk (bokmål)
- Polski
- Português
- Română
- Ρωσικά
- Simple English
- Slovenčina
- Svenska
- Türkçe
- Українська
- 中文

[[Επεξεργασία](#)]

Πίνακας περιχομένων ([Λεπτομέρεια](#))

- 1 [Διοικητική διαίρεση του δήμου](#)
- 2 [Περιγραφή](#)
- 3 [Εξωτερικοί σύνδεσμοι](#)
- 4 [Φωτογραφικό υλικό](#)

Διοικητική διαίρεση του δήμου [[Επεξεργασία](#)]

Ο δήμος περιλαμβάνει 5 δημοτικά διαμερίσματα:

- Κόρινθου (29.787 κάτοικοι)
- Αρχαία Κόρινθου (2.578 κάτοικοι)
- Εξαμίλιων (2.224 κάτοικοι)
- Ξυλοκέρκρας (1.215 κάτοικοι)
- Σαλιμού (751 κάτοικοι)

Περιγραφή [[Επεξεργασία](#)]

Η πόλη περιτοχιζείται από τις παραθαλάσσιες περιοχές του Λεχαιού, Ισθμίων, Κεραίων και τις περιοχές Εξαμίλιων και Αρχαίας Κόρινθου

Η γεωγραφική της θέση είναι στην παραθαλάσσια περιοχή ανατολικά του [Κορινθιακού κόλπου](#) κοντά στην ζεύξη με τον Σαρωνικό

Περιβάλλεται από το βουνό [Όνεια](#) και [Ακρακόρινθο](#), όπου κατά την αρχαιότητα βρισκόταν η αρχαία πόλη της Κόρινθου.


Η Κόρινθος βρίσκεται στον οδικό άξονα [Αθήνα](#) - [Πάτρα](#), με την πρόσβαση στην πρωτεύουσα να είναι εύκολη για τους κατοίκους της είτε μέσω της σύγχρονης εθνικής οδού είτε μέσω του [προαστιακού σιδηροδρόμου](#) που άρχισε να λειτουργεί το [2005](#).

Από το ακαδημαϊκό έτος [2003-2004](#) λειτουργεί στην Κόρινθο το [Τμήμα Κοινωνικής και Εκπαιδευτικής Πολιτικής](#) της Σχολής Κοινωνικών Επιστημών του [Πανεπιστημίου Πελοποννήσου](#)


Εξωτερικοί σύνδεσμοι [[Επεξεργασία](#)]

- [Ο σπάλυτος οδηγός για την Κόρινθο](#)
- <http://www.korinthos.gr>
- <http://www.e-korinthiakos.gr>
- <http://www.uop.gr/departments/depart8/index.shtml>
- [Φωτογραφίες Κόρινθου](#)
- [Ζωντανή εικόνα \(live webcam\) από τον ναό του Απόλλωνα](#)


Φωτογραφικό υλικό [[Επεξεργασία](#)]




[Άποψη της Κόρινθου από το σιδηροδρομικό σταθμό σε επιστολικό δελτάριο των αρχών του 20ου αι.](#)




[Ισθμός της Κόρινθου](#)




Η γέφυρα του Ισθμού της Κόρινθου σε επιστολικό δελτάριο των αρχών του 20ου αι.



ναός του Απόλλωνα στην Αρχαία Κόρινθο.





Ο τύραννος της Κόρινθου, Περικλής (περίπου 627-585 π.Χ).




Η ναός του Αγίου Παύλου στην Κόρινθο σε επιστολικό δελτάριο των αρχών του 20ου αι.

Κόρινθος






Γεωγραφικό διαμέρισμα	Πελοποννήσου
Περιφέρεια Νεώδος	Πελοποννήσου Κορινθίας
Αρ. Δευ. Δημοτεσιών	5
Επίσημος πληθυσμός	29.787 (2001)
Τελεφεφωνικός κώδικας	201 00
Τηλεφωνικός κώδικας	2741
Αιχμαλωτιστής	Αλέξανδρος Πνευματικός
Αιχμαλωτικός τόπος	http://www.korinthos.gr/



Τα [Κοινά](#) έχουν πολυμέσα σχετικά με το άρθρο: [Κόρινθος](#)

Σχήμα 6.29 - Ιστοσελίδα της Βικιπαίδειας για την Κόρινθο με την σωστή γεωκωδικοποίηση



ΒΙΚΙΠΑΙΔΕΙΑ
Η ελεύθερη εγκυκλοπαίδεια

πλοήγηση

- Κύρια Σελίδα
- Κατάλογος άρθρων
- Αλφάβητα άρθρα
- Τρέχοντα γεγονότα
- Τυχαία σελίδα

συμμετοχή

- Πρόσφατες αλλαγές
- Αγορά
- Πύλη Κοινότητας
- Βοήθεια
- Επικοινωνία
- Δικρές

αναζήτηση

Μετάβαση
Αναζήτηση

εργαλειοθήκη

- Συνδέσεις προς εδώ
- Σχετικές αλλαγές
- Επιμέλεια αρχείου
- Εδικές σελίδες
- Έκδοση εκτύπωσης
- Σφραγές συνδέσεων
- Βιβλιογραφικές πηγαίνοτες

άλλες γλώσσες

- English
- Italiano
- Lumbeeit

άρθρο συζήτηση επεξεργασία ιστορικό

Λεόντιο Αχαΐας

Από τη Βικιπαίδεια, την ελεύθερη εγκυκλοπαίδεια
(Ανακατεύθυνση από [Λέοντα](#))

Το **Λεόντιο** είναι ανεξάρτητη κοινότητα της ορεινής **Αχαΐας** ("Καποδιστριακή κοινότητα"). Μέχρι το 1914 ήταν δημοτικό διαμέρισμα του τότε **Δήμου Φαρρών** , και από τότε ανεξάρτητη κοινότητα. Βρίσκεται σε ύψος 750-800 μέτρων στους πρόποδες του **Παναχαϊκού**. Πλησίον του χωριού υπάρχουν πολλές πηγές με κρυστάλλινα, πεντακάθαρα νερά καθώς και ένας παρατάμαος του Ξελονούτα.

Πίνακας περιεχομένων [[Διεύθυνση](#)]

1 [Όνομασία](#)

2 [Ιστορία](#)

3 [Λαϊκές παραδόσεις](#)

4 [Αναφορές](#)

Όνομασία [[Επεξεργασία](#)]

Το σημερινό χωριό διατηρεί το όνομα της αρχαίας πόλης που υπήρχε εκεί.

Ιστορία [[Επεξεργασία](#)]

Διαβάστε για το αρχαίο Λεόντιο
Εδώ

Ανακαφέες που έγιναν κατά τη δεκαετία του 1930 έφεραν στο φως τάφους και ευρήματα της Μυκηναϊκής εποχής 14ος-12ος αι. π.Χ. Υπάρχουν επίσης ίχνη ακρόπολης στο εκκλησάκι του Αγίου Κωνίνου που είναι ορατά ως σήμερα γύρω στον 3ο αι π.Χ. Λίγο πιο κάτω εντοπίστηκαν τα θεμέλια πρωτοχριστιανικής βασιλικής μεγάλων διαστάσεων πάνω στο πλάτωμα των οποίων είναι κτισμένος ο ναός του Αγίου Ανδρέα (4ος-5ος αι μ.Χ). Αυτά τα στοιχεία υποδηλώνουν συνεχή και έντονη ανθρώπινη παρουσία στην περιοχή, δυστυχώς όμως η έντονη διάβρωση λόγω των πολλών υδάτων, η εγκατάλειψη και οι περίοδοι αναστατώσεων του μεσαιώνα λίγα πράγματα δέσωσαν από την αλλοτινή ακμή της περιοχής.

Το Λεόντιο είχε δημιουργήσει και αποικία ^[1] στην **Χίο** τα **Λεοντόποδα**.

Κατά τον **Πύλο** στο Λεόντιο υπήρχε μία πηγή ^[2] που όποιος έπινε νερό πέθανε μετά από τρεις ημέρες.

Στη σύγχρονη ιστορική σημαντικά γεγονόσ, που έλαβε χώρα στα περίγυρα του χωριού είναι η γνωστή **Μάχη του Λεοντίου** ανάμεσα σε ομάδα αναρτών του ΕΛΑΣ και του 3ου συντάγματος "εδικού πολέμου" της παλαιής Μεραρχίας Πεζομέτε υπό τον ταγματάρχη Γκασιάρδο, τον επονιομαζόμενο και "αναρτοεθγό" Επικεφαλής των αναρτών ήταν ο καπετάνιος Μήκος. Στις 14 Ιουλίου του 1943 ξεκίνησε η μάχη με ενόρα των αναρτών. Στη μάχη που ακολούθησε οι έλληνες πατριώτες κυριολεκτικά διέλυσαν τις ιταλικές κατοχικές δυνάμεις οι οποίες έχασαν γύρω στους εβδομήντα άνδρες όλο τον οπλισμό και τα υποζύγια τους. Με τα υλικά αυτά ενασχόθηκε αποασιαστικά η μαχητική ικανότητα του ΕΛΑΣ της βόρειας Πελοποννήσου. Αργότερα ως αντίποινα το χωριό κάηκε από τις κατοχικές δυνάμεις που σκότωσαν τουλάχιστον δέκα χωρικούς. Πολλοί κάτοικοι του Λεοντίου συμμετείχαν ενεργά στον εφεδρικό ΕΛΑΣ και τις οργανώσεις του ΕΑΜ παρέχοντας ουσιαστική υποστήριξη στον απελευθερωτικό αγώνα τα χρόνια της κατοχής 1941-1944.

Λαϊκές παραδόσεις [[Επεξεργασία](#)]

Απο το Λεόντιο πέρασε ο **Απόστολος Ανδρέας**^[3] αφού περιπλανήθηκε κυνηγημένος απο τους Ρωμαίους έφτασε στο χωριό. Αλλά οι κάτοικοι του τον αντιμετώπισαν άσχημα μην ξέροντας ποιάς είναι. Μόλις στην θέση **Πάτρα** υπάρχουν δύο πλάκες που λένε ότι εκεί πάτησε ο Άγιος , στην μία αποτυπώθηκε το παπούτσι του και στην άλλη το δεξί του ποδι αφού είχε χάσει το παπούτσι.

Μιά άλλη παράδοση λέει ^[4] ότι ο Απόστολος ερχόμενος εκεί κυνηγημένος, απο τους στρατιώτες που τον έψαχναν για να τον σταυρώσουν, συνάντησε έναν χωριάτη που που έπαιρνε το χωράφι του και του λέει "αν ένας στρατιώτης και σε ρωπήσουν, που ίσα κρύφτηκε , να μην το μαρτυρήσεις , πως ήρθα απο εδώ , γιατί θέλουν να με σταυρώσουν". Ο χωριάτης δεν ήξερε ποιά ήταν και απόρησε όταν είδε ότι μπροστά έπαιρνε και πίσω φύτρωνε, και πρόδωσε τον άγιο μαρτυρώντας τον στους στρατιώτες.Αυτοί τον έδεσαν πισθάγκωνα και τον πήγαν στην **Πάτρα** και τον σταυρώσατε.

Λεόντιο

ΝομόςΑχαΐας

Δίμοος**Κοινότητα Λεοντίου**

Υψόμετρο800 m

Πρώην όνομαΓουρζούμια

Αναφορές [[Επεξεργασία](#)]

- ↑ Αδαμάντιος Κοραΐς Άτακτα
- ↑ Πλίνιος Ν.HIST.XXXI.19
- ↑ *Στέφανος Ουμπόπουλος*, *Ιστορία της Πόλεως των Πατρών απο αρχαιοτάτων χρόνων μέχρι το 1821*, Τόμος Α', *Αχαϊκές Εκδόσεις*, *Πάτρα* 1998. *ISBN 980-7960-10-6*
- ↑ *Στέφανος Ουμπόπουλος*, *Ιστορία της Πόλεως των Πατρών απο αρχαιοτάτων χρόνων μέχρι το 1821*, Τόμος Α', *Αχαϊκές Εκδόσεις*, *Πάτρα* 1998. *ISBN 980-7960-10-6*

Δήμοι και κοινότητες του νομού [[απόκρυψη](#)]

Αχαΐας

Δήμος Πατρέων | Δήμος ~~Συμμετοχής~~ | Δήμος Αντιόχ | Δήμος

Ακρόπολις

Δήμος Αρκαονίας

Δήμος Βοηχολοίων | Δήμος

Διαμαντιού

Δήμος Διμης

Δήμος Ερινεύου

Δήμος Καλαβρύτων

Δήμος Λαρισσού

Δήμος Λευκασίου

Δήμος Μεσσήνης

Δήμος Μόβρης

Δήμος Πάων

Δήμος Περαλίας

Δήμος Ρίου

Δήμος Συμμετοχής

Δήμος Τριπλίας

Δήμος Φαρρών

Δήμος Ολυνθίας

Κοινότητα Καλεντζίου

Κοινότητα Λεοντίου

Κατηγορίες

- Κοινότητα [Λεοντίου](#)
- Χωριά της [Αχαΐας](#)

Τελευταία τροποποίηση 19:56, 12 Απριλίου 2008. Όλα τα κείμενα είναι διαθέσιμα υπό την [GNU Free Documentation License](#) (βλέπε [Βικιπαίδεια:Πνευματικά δικαιώματα για Απριλέρτες](#)).

[Προστασία Προσωπικών Δεδομένων](#)
[Για τη Βικιπαίδεια](#)
[Απαιτήσεις ευθυνών](#)

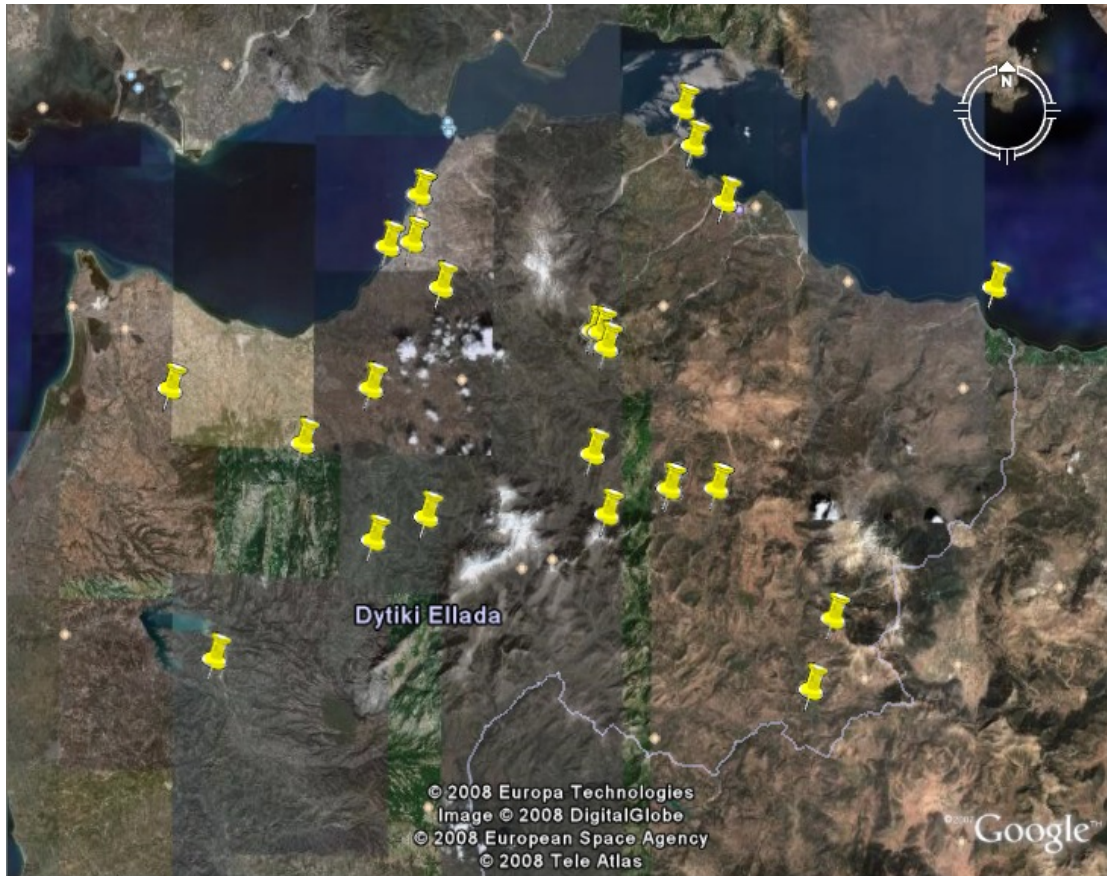
Σχήμα 6.30 - Ιστοσελίδα της Βικιπαίδειας για το Λεόντειο Κορινθίας με την σωστή γεωκωδικοποίηση

Συγκρίνοντας τα αποτελέσματα στα Σχήμα 6.29 και Σχήμα 6.30 με τα αντίστοιχα των 3.2 και 3.4 παρατηρούμε πως στα δύο πρώτα σχήματα υπάρχει μια ολοκληρωμένη απεικόνιση των γεωγραφικών πληροφοριών των ιστοσελίδων αυτών.

Στην συνέχεια θα παρουσιάσουμε με την βοήθεια του Google Earth τα διορθωμένα αποτελέσματα συνολικά σε ένα χάρτη της Ελλάδος.

Σχήμα 6.31- Βελτιωμένη συνολική απεικόνιση της ιστοσελίδας για την Κόρινθο





Σχήμα 6.32 - Βελτιωμένη συνολική απεικόνιση της ιστοσελίδας για το Λεόντειο

Όπως βλέπουμε από τους δύο προηγούμενους χάρτες δεν υπάρχουν σημεία σε περιοχές που δεν περιέχονται στο κείμενο, κάτι που συνέβαινε στα Σχήμα 3.7 και Σχήμα 3.9.

7

Επίλογος

7.1 Σύνοψη και συμπεράσματα

Στις προηγούμενες ενότητες παρουσιάστηκε ένα ολοκληρωμένο εργαλείο για την απεικόνιση των γεωγραφικών πληροφοριών μιας ιστοσελίδας την ώρα που είναι ανοιχτή σε έναν φυλλομετρητή Mozilla FireFox. Εκτός όμως την απεικόνιση των πληροφοριών το εργαλείο αυτό προσφέρει την δυνατότητα στον χρήστη να βελτιώσει τα αποτελέσματα που εμφανίζονται εκείνη την ώρα στην οθόνη με την βοήθεια ενός εύχρηστου menu.

Το εργαλείο αυτό τελικά δημιουργεί μια ενιαία βάση δεδομένων με πληροφορίες από όλες τις ιστοσελίδες στις οποίες έχει χρησιμοποιηθεί το εργαλείο, κάτι που μπορεί να βοηθήσει στην δημιουργία γεωγραφικών λεξικών. Το γεγονός ότι διανέμεται με την μορφή ενός extension, το κάνει αρκετά λειτουργικό, αποδοτικό και εύκολο στην εγκατάσταση.

Αυτή η δυνατότητα του χρήστη να βελτιώνει τα όποια λάθη βλέπει, κρίθηκε απαραίτητη αφού οι αλγόριθμοι αυτόματης γεωκωδικοποίησης δεν είναι αρκετοί. Παρόλα τη θετική συνεισφορά που έχει αυτή η πράξη, η βελτίωση των αποτελεσμάτων δεν παύει να εξαρτάται από τις γεωγραφικές γνώσεις του χρήστη και τον εθελοντισμό που μπορεί να υπάρξει.

7.2 Μελλοντικές επεκτάσεις

Στην ενότητα αυτή θα παρουσιάσουμε κάποιες ιδέες για την επέκταση της παρούσας διπλωματικής.

7.2.1 Δημιουργία μίας δημόσιας αποθήκης ή υπηρεσίας για την

γεωκωδικοποίηση ιστοσελίδων

Όλα όσα περιγράφηκαν στις προηγούμενες ενότητες κάνουν ιδιαίτερα σαφές το γεγονός πως η συγκεκριμένη διπλωματική εργασία για να εφαρμοστεί και προσφέρει όλα τα πλεονεκτήματα, στα οποία εκτενώς αναφερθήκαμε, θα πρέπει να μπορεί να χρησιμοποιηθεί από κάθε χρήστη του διαδικτύου χωρίς να του δημιουργεί προβλήματα. Έτσι λοιπόν απαιτείται το τμήμα της διπλωματικής που αφορά το extension να είναι σε μορφή που να προσφέρει αυτόματη εγκατάσταση του extension σε φυλλομετρητή Mozilla FireFox καθώς επίσης η κεντρική βάση δεδομένων να είναι εγκατεστημένη σε κάποιον δημόσιο server ώστε να δημιουργηθεί ένα είδος δημόσιας αποθήκης όπου τα αποθηκεύονται τα αποτελέσματα της αυτόματης και χειρονακτική γεωκωδικοποίησης που προκαλεί ο κάθε χρήστης ξεχωριστά. Τα αποτελέσματα που συσσωρεύονται θα πρέπει να ελέγχονται για λόγους κυρίως αξιοπιστίας και εγκυρότητας.

7.2.2 Λογισμικό Ανοιχτού Κώδικά

Χρησιμοποιώντας εργαλεία που ανήκουν στην συνομοταξία των εφαρμογών ανοιχτού κώδικα είναι αναπόφευκτο για την παρούσα διπλωματική εργασία να αποτελεί επίσης εφαρμογή ανοιχτού κώδικα. Όμως εκτός από τον ηθικό αυτό λόγο, όλη η φιλοσοφία πάνω στην οποία στηρίζεται η εφαρμογή μας οδηγεί τελικά στον ανοιχτό κώδικα. Έτσι θα μπορούν να προσφέρονται δύο από τα πιο βασικά πλεονεκτήματα των εφαρμογών ανοιχτού κώδικα, η δυνατότητα πρόσβασης στον πηγαίο κώδικα της εφαρμογής και τροποποίηση του από όλους. Εκτός από την δημοσιοποίηση του κώδικα για να χαρακτηριστεί «ανοικτού κώδικα» ένα τέτοιο προϊόν θα πρέπει να χαρακτηρίζεται από λειτουργικότητα δηλ. να μπορεί να λειτουργήσει σε οποιαδήποτε πλατφόρμα, στην περίπτωση μας να μπορεί να ενσωματωθεί εύκολα σε οποιοδήποτε φυλλομετρητή (π.χ. Internet Explorer) σε όλα τα λειτουργικά συστήματα (π.χ. Linux).

8

Βιβλιογραφία

- [1] McCurley, K. Geospatial mapping and navigation of the web. In Proc. 10th international conference on World Wide Web, pp. 221-229, 2001.
- [2] Explore Our Pla.Net. RSS to GeoRSS converter. Ιστοσελίδα <http://exploreourpla.net/2006-06-08/georss-feed-reader-shows-podcasts.html> , 2006.
- [3] Luis Gravano, Vasileios Hatzivassiloglou, Richard Lichtenstein. Categorizing web queries according to geographical locality. In Proc. of CIKM, pp. 325-333, 2003.
- [4] Junyan Ding, Luis Gravano, Narayanan Shivakumar. Computing Geographical Scopes of Web Resources. In Proc. of VLDB, pp 545-556, 2000.
- [5] Karla A. V. Borges, Alberto H. F. Laender, Claudia Bauzer Medeiros , Altigran, The Web as a Data Source for Spatial Databases, GeoInfo, 2003
- [6] Einat Amitay, Nadav Har'EL, Ron Sivan, Aya Soffer. Web-a-Where: Geotagging Web Content. In Proc. of SIGIR, pp. 273-280, 2004.
- [7] Amittai E. Axelrod. On Building a High Performance Gazetteer Database. MetaCarta, Inc.
- [8] Otis Gospodnetic, Erik Hatcher, Lucene in Action, Manning, page 138, 2005
- [9] E. Brill, A Simple Rule-based Part of Speech Tagger, In Proc. of Third Conference on Applied Natural Language Processing, 1992
- [10] Rahm E., Do H.H, Data Cleaning: Problems and Current Approaches, In Proc. Of IEEE Bulletin on Data Engineering, vol 23(4), 2000

- [11] <http://www.networkingfiles.com/neotrace/>. Ίσχυε τον Μάιο του 2008.
- [12] <http://www.metacarta.com>. Ίσχυε τον Μάιο του 2008.
- [13] Free Software Foundation, Inc. Wikipedia – Mozilla FireFox. Ιστοσελίδα http://el.wikipedia.org/wiki/Mozilla_Firefox. Ίσχυε τον Μάιο του 2008.
- [14] Mozilla Europe and Mozilla Foundation. Mozilla FireFox. Ιστοσελίδα <http://www.mozilla-europe.org/el/products/firefox/>. Ίσχυε τον Μάιο 2008.
- [15] Free Software Foundation, Inc. Wikipedia - Mozilla Extension. Ιστοσελίδα http://en.wikipedia.org/wiki/Extension_%28Mozilla%29. Ίσχυε τον Μάιο του 2008.
- [16] Individual mozilla.org contributors. Mozilla Developer Center - Extensions. Ιστοσελίδα <http://developer.mozilla.org/en/docs/Extensions>. Ίσχυε τον Μάιο του 2008.
- [17] Individual mozilla.org contributors. Mozilla Developer Center – Set up extension Enviroment. Ιστοσελίδα http://developer.mozilla.org/en/docs/Setting_up_extension_development_environment. Ίσχυε τον Μάιο του 2008.
- [18] roachfiend.com. How to create a Firefox Extension. Ιστοσελίδα <http://roachfiend.com/archives/2004/12/08/how-to-create-FireFox-extensions>. Ίσχυε τον Μάιο του 2008.
- [19] MozillaZine. Getting started with extension development. Ιστοσελίδα http://kb.mozillazine.org/Getting_started_with_extension_development. Ίσχυε τον Μάιο του 2008.
- [20] Mozilla.org. Tutorial : Creating a Mozilla Extension. Ιστοσελίδα <http://www.mozilla.org/docs/tutorials/tinderstatus>. Ίσχυε τον Μάιο του 2008.
- [21] Jonah Bishop. Firefox Toolbar Tutorial. Ιστοσελίδα <http://www.borngeek.com/firefox/toolbar-tutorial/>. Ίσχυε τον Μάιο του 2008.
- [22] Tippit, Inc. SoftwareDeveloper. Ιστοσελίδα <http://www.softwaredeveloper.com/features/firefox-extension-resource-072307/>. Ίσχυε τον Μάιο του 2008.
- [23] Individual mozilla.org contributors. Code snippets. Ιστοσελίδα http://developer.mozilla.org/en/docs/Code_snippets. Ίσχυε τον Μάιο του 2008.
- [24] MozillaZine. Example_code. Ιστοσελίδα http://kb.mozillazine.org/Category:Example_code. Ίσχυε τον Μάιο του 2008.
- [25] MozillaZine. Forum. Ιστοσελίδα <http://forums.mozillazine.org/viewforum.php?f=27>. Ίσχυε τον Μάιο του 2008.
- [26] Individual mozilla.org contributors. Chrome Manifest. Ιστοσελίδα http://developer.mozilla.org/en/docs/Chrome_Registration. Ίσχυε τον Μάιο του 2008.

- [27] MozillaZine. Install.rdf. Ιστοσελίδα <http://kb.mozillazine.org/Install.rdf>. Ίσχυε τον Μάιο του 2008.
- [28] MozillaZine. Profile Manager. Ιστοσελίδα http://kb.mozillazine.org/Profile_manager. Ίσχυε τον Μάιο του 2008.
- [29] Free Software Foundation, Inc. Wikipedia. Ιστοσελίδα http://en.wikipedia.org/wiki/Ajax_%28programming%29. Ίσχυε τον Μάιο του 2008.
- [30] Refsnes Data. Ajax tutorial. Ιστοσελίδα <http://www.w3schools.com/Ajax/Default.Asp>. Ίσχυε τον Μάιο του 2008.
- [31] Individual mozilla.org contributors. Ajax : Getting Started. Ιστοσελίδα http://developer.mozilla.org/en/docs/AJAX:Getting_Started. Ίσχυε τον Μάιο του 2008.
- [32] O'Reilly Media, Inc. A simpler Ajax Path. Ιστοσελίδα <http://www.onlamp.com/pub/a/onlamp/2005/05/19/xmlhttprequest.html>. Ίσχυε τον Μάιο του 2008.
- [33] Jupitermedia, Org. Alternate Java Techniques. Ιστοσελίδα http://www.webreference.com/programming/ajax_tech/. Ίσχυε τον Μάιο του 2008.
- [34] Adaptive Path Inc. Ajax : A New Approach to Web Applications. Ιστοσελίδα <http://www.adaptivepath.com/publications/essays/archives/000385.php>. Ίσχυε τον Μάιο του 2008.
- [35] W3C. XMLHttpRequest. Ιστοσελίδα <http://www.w3.org/TR/XMLHttpRequest/>. Ίσχυε τον Μάιο του 2008
- [36] Sun Microsystems, Inc. Asynchronous JavaScript Technology and XML (Ajax) With the Java Platform. Ιστοσελίδα <http://java.sun.com/developer/technicalArticles/J2EE/AJAX/>. Ίσχυε τον Μάιο του 2008.
- [37] MozillaZine. Dom Inspector. Ιστοσελίδα http://kb.mozillazine.org/DOM_Inspector. Ίσχυε τον Μάιο του 2008.
- [38] W3C. DOM. Ιστοσελίδα <http://www.w3.org/DOM/>. Ίσχυε τον Μάιο του 2008.
- [39] Google Maps
- [40] Free Software Foundation, Inc. Wikipedia. Ιστοσελίδα http://en.wikipedia.org/wiki/Google_maps. Ίσχυε τον Μάιο του 2008.
- [41] Google Inc. Google Maps. Ιστοσελίδα <http://maps.google.com/>. Ίσχυε τον Μάιο του 2008.
- [42] Google Inc. Google Maps API. Ιστοσελίδα <http://code.google.com/apis/maps/>. Ίσχυε τον Μάιο του 2008.
- [43] Google Inc. Google Maps API Documentation. Ιστοσελίδα <http://code.google.com/apis/maps/documentation/>. Ίσχυε τον Μάιο του 2008.

- [44] Google Inc. Google Maps Help Center. Ιστοσελίδα <http://maps.google.com/support/>. Ίσχυε τον Μάιο του 2008.
- [45] Google Inc. Google local. Ιστοσελίδα <http://local.google.com>. Ίσχυε τον Μάιο του 2008
- [46] Google Inc. Google Earth. Ιστοσελίδα <http://earth.google.com>. Ίσχυε τον Μάιο του 2008
- [47] Free Software Foundation, Inc. Wikipedia. Ιστοσελίδα <http://el.wikipedia.org/wiki/JavaScript>. Ίσχυε τον Μάιο του 2008.
- [48] Individual mozilla.org contributors. Mozilla Developer Center - JavaScript. Ιστοσελίδα <http://developer.mozilla.org/en/docs/JavaScript>. Ίσχυε τον Μάιο του 2008.
- [49] Refsnes Data. W3schools JavaScript Tutorial. Ιστοσελίδα http://www.w3schools.com/js/js_intro.asp. Ίσχυε τον Μάιο του 2008.
- [50] Free Software Foundation, Inc. Wikipedia. Ιστοσελίδα <http://en.wikipedia.org/wiki/Postgres>. Ίσχυε τον Μάιο του 2008.
- [51] PostgreSQL Global Development Group. PostgreSQL. Ιστοσελίδα <http://www.postgresql.org/>. Ίσχυε τον Μάιο του 2008.
- [52] Tomcat
- [53] Free Software Foundation, Inc. Wikipedia. Ιστοσελίδα http://en.wikipedia.org/wiki/Apache_Tomcat. Ίσχυε τον Μάιο του 2008.
- [54] The Apache Software Foundation. Apache Tomcat. Ιστοσελίδα <http://tomcat.apache.org/>. Ίσχυε τον Μάιο του 2008.
- [55] Servlets
- [56] Free Software Foundation, Inc. Wikipedia. Ιστοσελίδα <http://en.wikipedia.org/wiki/Servlet>. Ίσχυε τον Μάιο του 2008.
- [57] Sun Microsystems, Inc. Java Servlet Technology. Ιστοσελίδα <http://java.sun.com/products/servlet/>. Ίσχυε τον Μάιο του 2008.
- [58] XULPlanet.com. XULPlanet .Ιστοσελίδα <http://www.xulplanet.com/>. Ίσχυε τον Μάιο του 2008.
- [59] Individual mozilla.org contributors. Mozilla Developer Center - XUL. Ιστοσελίδα <http://developer.mozilla.org/en/docs/XUL>. Ίσχυε τον Μάιο του 2008.
- [60] XULPlanet.com. Forum. Ιστοσελίδα <http://www.xulplanet.com/forum2/>. Ίσχυε τον Μάιο του 2008.
- [61] XULPlanet.com. Overlay.xul .Ιστοσελίδα <http://xulplanet.com/tutorials/xultu/overlay.html>. Ίσχυε τον Μάιο του 2008.

- [62] Individual mozilla.org contributors. Mozilla Developer Center –Overlay.xul. Ιστοσελίδα http://developer.mozilla.org/en/docs/XUL_Overlays. Ίσχυε τον Μάιο του 2008.
- [63] Free Software Foundation, Inc. Wikipedia. Ιστοσελίδα http://el.wikipedia.org/wiki/Java#.CE.A4.CE.B1_.CF.87.CE.B1.CF.81.CE.B1.CE.BA.CF.84.CE.B7.CF.81.CE.B9.CF.83.CF.84.CE.B9.CE.BA.CE.AC_.CF.84.CE.B7.CF.82_Java. Ίσχυε τον Μάιο του 2008.
- [64] Free Software Foundation, Inc. Wikipedia. Ιστοσελίδα <http://en.wikipedia.org/wiki/Java>. Ίσχυε τον Μάιο του 2008.
- [65] Sun Microsystems, Inc. The Source for Java Developers. Ιστοσελίδα <http://java.sun.com/>. Ίσχυε τον Μάιο του 2008.
- [66] Free Software Foundation, Inc. Wikipedia. Ιστοσελίδα <http://en.wikipedia.org/wiki/Php>. Ίσχυε τον Μάιο του 2008.
- [67] The PHP Group. PHP Documentation. Ιστοσελίδα <http://www.php.net/manual/el/preface.php>. Ίσχυε τον Μάιο του 2008.
- [68] The Apache Software Foundation. The Apache HTTP Server Project. Ιστοσελίδα <http://httpd.apache.org/>. Ίσχυε τον Μάιο του 2008.
- [69] Clearforest. Clearforest Gnosis extension. Ιστοσελίδα http://sws.clearforest.com/Blog/?page_id=32/. Ίσχυε τον Μάιο του 2008.
- [70] Ben Basson. Context Highlight 3.0 extension. Ιστοσελίδα <http://www.cusser.net/extensions/contexthighlight/>. Ίσχυε τον Μάιο του 2008.
- [71] Jeremy Gillick. TinyUrl creator extension. Ιστοσελίδα <https://addons.mozilla.org/en-US/firefox/addon/126>. Ίσχυε τον Μάιο του 2008.
- [72] Mozdev.org. Dictionary Search extension. Ιστοσελίδα <http://dictionarysearch.mozdev.org/>. Ίσχυε τον Μάιο του 2008.
- [73] Masato Shimoyama. MSTTS Dictionary extension. Ιστοσελίδα <http://homepage1.nifty.com/masato74/msts-dict/>. Ίσχυε τον Μάιο του 2008.
- [74] The Eclipse Foundation. Eclipse.org Home. Ιστοσελίδα <http://www.eclipse.org/>. Ίσχυε τον Μάιο του 2008.
- [75] Αλβέρτου-Δαυίδ Αντζελ . Διπλωματική “Έξαγωγή Γεωγραφική Πληροφορίας από Ημιδομημένο Κείμενο”. Ιούλιος 2006.
- [76] Nikolas C.Zakas, Jeremy McPeak and Joe Fawcett. Professional AJAX. Wrox Press 2006.