



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

Geotyping
Live γεωκωδικοποίηση ελεύθερου κειμένου

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

ΒΙΔΡΟΥ ΣΩΚΡΑΤΗ

Επιβλέπων : Τίμος Σελλής
Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2011



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

Geotyping

Live γεωκωδικοποίηση ελεύθερου κειμένου

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

ΒΙΑΡΟΥ ΣΩΚΡΑΤΗ

Επιβλέπων : Τίμος Σελλής
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 4^η Ιουλίου 2011.

.....
Τ. Σελλής
Καθηγητής Ε.Μ.Π.

.....
Ι. Βασιλείου
Καθηγητής Ε.Μ.Π.

.....
Dieter Pfoser
Ερευνητής Β' ΠΣΥΠ/Ε.Κ.
"Αθηνά"

Αθήνα, Ιούλιος 2011

.....
ΒΙΔΡΟΣ ΣΩΚΡΑΤΗΣ

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

Copyright © Σωκράτης Κ. Βίδρος, 2011.

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

Τα τελευταία χρόνια υπάρχουν σημαντικές εξελίξεις τόσο στο θέμα της γεωκωδικοποίησης¹, όσο και στον τομέα του geoparsing². Όλες σχεδόν οι υπηρεσίες χαρτών όπως είναι το Google Maps, το Yahoo Maps και το Open Street Maps παρέχουν υπηρεσίες geocoding, με διαφορετικούς βαθμούς ακρίβειας. Πρόοδος έχει γίνει και στο κομμάτι του geoparsing, αφού υπάρχουν εργαλεία, όπως το Yahoo Placemaker³ που προσπαθούν να συνδυάσουν τις παραπάνω τεχνικές.

Το ΠΣΥΠ έχει σημαντική εμπειρία στο συγκεκριμένο τομέα, καθώς έχει γίνει σημαντική δουλειά πάνω στην αναγνώριση γεωγραφικών όρων σε ελληνικό ελεύθερο κείμενο, ενώ υπάρχει σημαντικό know how στη χρήση των υπηρεσιών γεωκωδικοποίησης.

Στόχος της παρούσης διπλωματικής εργασίας είναι η ανάπτυξη μιας AJAX web εφαρμογής η οποία θα χρησιμοποιεί τις ήδη υπάρχουσες δυνατότητες αναγνώρισης γεωγραφικών όρων σε ελληνικό και αγγλικό ελεύθερο κείμενο από προηγούμενα εργαλεία σε συνδυασμό με online υπηρεσίες γεωκωδικοποίησης, προκειμένου να εντοπίζονται οι γεωγραφικοί όροι που πληκτρολογεί ο χρήστης και να είναι δυνατή η ταυτόχρονη απεικόνισή τους στον χάρτη. Η εφαρμογή θα ενσωματωθεί σε ένα σύστημα διαχείρισης κειμένου.

Λέξεις Κλειδιά: Geotyping, Geoparsing, Γεωκωδικοποίηση, Ajax, Google Maps, CKEditor, Εξαγωγή γεωγραφικής πληροφορίας, Σύστημα διαχείρισης κειμένου

¹ αντιστοίχιση κειμένου σε συντεταγμένες

² αναγνώριση γεωγραφικών όρων σε κείμενο

³ <http://developer.yahoo.com/geo/placemaker/>

Abstract

In recent years there has been a significant development in terms of geocoding⁴ and geoparsing⁵. Almost all online web mapping services like Google Maps, Yahoo Maps and Open Street Maps provide geocoding services, with varying degrees of accuracy. Progress has also been made to geoparsing techniques, since there are tools such as Yahoo Placemaker⁶, trying to perform geoparsing and geocoding simultaneously.

IMIS has a considerable experience on this field. The institute has carried out important research in the identification of geographic terms in Greek and English text, while there is considerable know how on geocoding services.

The scope of this thesis is to develop an AJAX web application that will use the existing possibilities for recognition of geographical terms in Greek and English text in conjunction with online geocoding services, in order to identify geographical terms in free text and enable simultaneous display on the map. The application will be embedded in a text management system.

Keywords: Geotyping, Geoparsing, Geocoding, Ajax, Google Maps, CKEditor, Geographical Information Extraction, Text Management Systems

⁴ matching text with coordinates

⁵ recognition of geographical terms in the text

⁶ <http://developer.yahoo.com/geo/placemaker/>

Ευχαριστίες

Αρχικά θα ήθελα να ευχαριστήσω τον καθηγητή κ. Τιμολέοντα Σελλή για την επίβλεψη αυτής της διπλωματικής εργασίας και κυρίως για την ευκαιρία που μου έδωσε να συνεργαστώ με το Ινστιτούτο Πληροφοριακών Συστημάτων και Προσομοίωσης.

Ευχαριστώ ιδιαίτερα τον Δρ. Dieter Pfoser για την καθοδήγηση του και την εξαιρετική συνεργασία που μου προσέφερε. Ακόμη, ευχαριστώ τους Ευθύμη Δρυμόνα και Αλέξανδρο Εφεντάκη για την πολύτιμη βοήθεια.

Τέλος θα ήθελα να ευχαριστήσω όλους τους καθηγητές και την οικογένειά μου για την καθοδήγηση και την ηθική συμπαράσταση που μου προσέφεραν όλα αυτά τα χρόνια ως φοιτητής στο Εθνικό Μετσόβιο Πολυτεχνείο.

Πίνακας περιεχομένων

1.	Εισαγωγή.....	1
1.1	Εξαγωγή γεωγραφικής πληροφορίας.....	1
1.2	Αντικείμενο διπλωματικής.....	2
1.2.1	Συνεισφορά.....	2
1.3	Οργάνωση κειμένου.....	3
2.	Σχετικές εργασίες.....	5
2.1	Ορισμοί.....	5
2.1.1	<i>Geoparsing</i>	5
2.1.2	<i>Γεωκωδικοποίηση</i>	6
2.1.3	<i>Geotagging</i>	6
2.1.4	<i>Geoblogging</i>	7
2.1.5	<i>Ajax</i>	7
2.2	Σχετικές Εργασίες.....	8
2.2.1	<i>Google Maps</i>	8
2.2.2	<i>Yahoo Placemaker</i>	9
2.2.3	<i>Geocrowd</i>	9
2.2.4	<i>Γεωκωδικοποίηση ιστοσελίδων</i>	9
2.2.5	<i>CKEditor</i>	10
2.2.6	<i>jQuery</i>	11
2.2.7	<i>Σύστημα διαχείρισης κειμένου</i>	11
3.	Ανάλυση Απαιτήσεων Συστήματος.....	13
3.1	Αρχιτεκτονική.....	13
3.1.1	<i>Geotyping</i>	13
3.1.2	<i>Περιγραφή Λειτουργιών</i>	15
3.1.2.1	<i>Geotyping Core</i>	15
3.1.2.2	<i>Google Maps</i>	17
3.1.2.3	<i>CKEditor</i>	18

3.1.2.4	Manual Tool.....	18
3.1.2.5	Geoparser	18
3.1.2.6	Εσωτερικός parser.....	19
3.1.2.7	Geotyping Hash.....	19
3.1.2.8	Highlighter	19
3.1.2.9	Proxy Server.....	19
3.1.2.10	Export Manager	19
3.1.2.11	Geotyping Plug-in for CKEditor.....	20
3.1.3	<i>Geotyping Text Management System</i>	20
3.1.4	<i>Περιγραφή λειτουργιών</i>	20
3.1.4.1	Διαχείριση χρηστών	20
3.1.4.2	Δημοσίευση και επεξεργασία άρθρων.....	21
3.1.4.3	Αναζήτηση άρθρων	21
3.1.4.4	Προβολή άρθρων	21
3.2	Σενάρια χρήσης.....	22
3.2.1	<i>Περίπτωση χρήσης "Δημοσίευση νέου άρθρου"</i>	24
3.3	Μοντέλο Οντοτήτων Συσχετίσεων	25
3.3.1	<i>Οντότητες</i>	25
3.3.1.1	Χρήστης	25
3.3.1.2	Ομάδα	25
3.3.1.3	Άρθρο.....	25
3.3.2	<i>Συσχετίσεις</i>	25
4.	Σχεδίαση Συστήματος Geotyping	27
4.1	Αρχιτεκτονική.....	27
4.2	Περιγραφή Κλάσεων	29
4.2.1	<i>Geotyping.js</i>	29
4.2.2	<i>Googlemaps.js</i>	31
4.2.3	<i>Gtm.js</i>	32
4.2.4	<i>Hash.js</i>	32
4.2.5	<i>Highlight.js</i>	32
4.2.6	<i>Parser.js</i>	33
4.2.7	<i>XmlWrite.js</i>	33

5.	Σχεδίαση συστήματος Geotyping Text Management System	35
5.1	Αρχιτεκτονική.....	35
5.1.1	<i>Model</i>	36
5.1.2	<i>View</i>	36
5.1.3	<i>Controller</i>	37
5.2	Οργάνωση φακέλων και αρχείων	38
5.3	Βασικά χαρακτηριστικά.....	39
5.4	Περιγραφή κλάσεων	40
5.4.1	<i>Γενικές κλάσεις εφαρμογής</i>	40
5.4.1.1	registry.class.php.....	40
5.4.1.2	router.class.php	40
5.4.1.3	template.class.php	41
5.4.2	<i>Controllers</i>	42
5.4.2.1	controller_base.class.php.....	42
5.4.2.2	adminController.php	42
5.4.2.3	articleController.php	43
5.4.2.4	editorController.php	43
5.4.2.5	error404.php	43
5.4.2.6	geotermsController.php.....	43
5.4.2.7	indexController.php.....	43
5.4.2.8	searchController.php	44
5.4.3	<i>Models</i>	45
5.4.3.1	article.class.php.....	45
5.4.3.2	articleManager.class.php.....	46
5.4.3.3	db.class.php.....	46
5.4.3.4	loggedInUser.class.php	46
5.4.3.5	mailer.class.php.....	46
5.4.3.6	newUser.class.php.....	46
5.4.3.7	userManager.class.php	46
5.4.4	<i>Views</i>	46
5.4.4.1	editor/account.php.....	47
5.4.4.2	editor/index.php	47
5.4.4.3	editor/post.php.....	47
5.4.4.4	contact.php	47

5.4.4.5	error404.php	47
5.4.4.6	forgot_password.php	47
5.4.4.7	index.php	47
5.4.4.8	listing.php	47
5.4.4.9	login.php	47
5.4.4.10	register.php	47
5.4.5	<i>lib</i>	48
5.4.5.1	funcs.general.php	48
5.4.5.2	funcs.user.php	48
6.	Υλοποίηση	49
6.1	Λεπτομέρειες υλοποίησης της εφαρμογής Geotyping	49
6.1.1	<i>Επισήμανση στοιχείων και Pipeline επεξεργασίας</i>	49
6.1.2	<i>Proxy Server και Cross-Domain Ajax Request</i>	51
6.1.3	<i>Google Bulk Geocoding</i>	51
6.1.4	<i>Ασάφεια κατά την γεωκωδικοποίηση</i>	52
6.1.5	<i>Υποστήριξη εικόνων</i>	52
6.1.6	<i>Μαύρη λίστα</i>	53
6.2	Λεπτομέρειες υλοποίησης του συστήματος Geotyping Text Management System .	54
6.2.1	<i>Registry</i>	54
6.2.2	<i>Διαβιβαστής και Search Engine Friendly Urls</i>	54
6.3	Πλατφόρμες και προγραμματιστικά εργαλεία	55
6.3.1	<i>Apache HTTP Server</i>	55
6.3.2	<i>MySQL</i>	56
6.3.3	<i>PHP</i>	57
6.3.4	<i>JavaScript, Βιβλιοθήκες και APIs</i>	57
6.3.5	<i>Πλατφόρμα ανάπτυξης</i>	58
7.	Έλεγχος και Εγκατάσταση	59
7.1	Εγκατάσταση	59
7.1.1	<i>Απαιτήσεις</i>	59
7.1.2	<i>Οδηγίες εγκατάστασης</i>	59
7.1.3	<i>Εγκατάσταση Geotyping</i>	60

7.2	Browser Compatibility	60
7.3	Εκκίνηση και χρήση της εφαρμογής.....	61
8.	Επίλογος	71
8.1	Σύνοψη και συμπεράσματα.....	71
8.2	Μελλοντικές επεκτάσεις	72
9.	Βιβλιογραφία.....	73

Κατάλογος σχημάτων

2-1 Η υπηρεσία Google Maps	8
2-2 Γραφικό περιβάλλον του CKEditor	10
3-1 Διάγραμμα υποσυστημάτων.....	15
3-2 Ακολουθιακό διάγραμμα Geotyping.....	17
3-3 Διάγραμμα χρήσης.....	23
3-4 Διάγραμμα Οντοτήτων-Συσχετίσεων.....	26
4-1 Προσχέδιο γραφικού περιβάλλοντος	27
5-1 Μοντέλο MVC με PHP και Apache Server	35
5-2 Ακολουθιακό διάγραμμα MVC προτύπου	36
5-3 Διάγραμμα γενικών κλάσεων εφαρμογής	40
5-4 Διάγραμμα κλάσεων ελεγκτών	42
5-5 Διάγραμμα κλάσεων μοντέλου	45
6-1 Geotyping Pipeline.....	49
6-2 Διάγραμμα καταστάσεων μηχανισμού μαρκαρίσματος.....	50
7-1 Αρχική σελίδα Geotyping Text Management System	61
7-2 Λίστα δημοσιευμένων άρθρων	62
7-3 Φόρμα επικοινωνίας.....	62
7-4 Σελίδα εγγραφής συντάκτη	63
7-5 Σελίδα ανάκτησης κωδικού.....	63
7-6 Σελίδα σύνδεσης συντάκτη	63
7-7 Αρχική σελίδα συντάκτη.....	64
7-8 Ρυθμίσεις λογαριασμού.....	64
7-9 Δημοσίευση νέου άρθρου	65
7-10 Γραφικό περιβάλλον του Geotyping	66
7-11 Τα μενού της εφαρμογής.....	66
7-12 Διαγραφή όρου.....	66

7-13 Χειροκίνητη γεωκωδικοποίηση φράσης	68
7-14 Διόρθωση γεωγραφικού όρου στον χάρτη	69
7-15 Υποστήριξη μορφοποιημένου κειμένου και εικόνων	70
7-16 Geotyping Plug-in for CKEditor	70

1. Εισαγωγή

1.1 Εξαγωγή γεωγραφικής πληροφορίας

Η ανάπτυξη της τεχνολογίας Ajax⁷ έδωσε την δυνατότητα δημιουργίας νέων καινοτόμων υπηρεσιών που επέτρεπαν την μεταφορά δεδομένων από και προς τον web server, χωρίς να απαιτείται η πλήρης επαναφόρτωση της ιστοσελίδας, μέσω ασύγχρονων HTTP αιτημάτων. Οι υπηρεσίες Google Maps και Gmail της Google είναι ίσως τα πιο γνωστά παραδείγματα τέτοιων υπηρεσιών. Παράλληλα η ενσωμάτωση γεωγραφικής πληροφορίας σε ολοένα και περισσότερες εφαρμογές καθώς επίσης και η ευρεία διαθεσιμότητα υπηρεσιών γεωκωδικοποίησης προσέφεραν νέες δυνατότητες επεξεργασίας και απεικόνισης πληροφοριών σε χάρτες.

Τα τελευταία χρόνια έχει σημειωθεί σημαντική πρόοδος στον τομέα την φυσικής επεξεργασίας γλώσσας. Έχουμε πλέον την δυνατότητα εξαγωγής γεωγραφικής πληροφορίας από μη δομημένο κείμενο με στόχο την απομόνωση γεωγραφικών όρων που μπορούν στην συνέχεια να γεωκωδικοποιηθούν. Το κείμενο μπορεί να είναι αποθηκευμένο σε κάποιο μέσο ή να εισάγεται από τον χρήστη μέσω των WYSIWYG⁸ επεξεργαστών κειμένου που προσφέρουν ολοκληρωμένη επεξεργασία κειμένου στον ιστό.

Όλα τα παραπάνω οδήγησαν στην ιδέα της ανάπτυξης μιας πλήρους δικτυακής εφαρμογής όπου ο χρήστης πληκτρολογεί το κείμενο του ενώ το σύστημα αναλαμβάνει, σε πραγματικό χρόνο, την εξαγωγή και την επεξεργασία της γεωγραφικής πληροφορίας που εντοπίζεται σε αυτό.

⁷ Asynchronous JavaScript

⁸ *What You See Is What You Get*. Ο όρος χρησιμοποιείται για να δηλώσει ένα σύστημα όπου το περιεχόμενο υπό επεξεργασία είναι όμοιο με το τελικό.

1.2 Αντικείμενο διπλωματικής

Αξιοποιώντας τις τεχνολογίες και τις υπηρεσίες που προαναφέρθηκαν, δημιουργήσαμε την εφαρμογή Geotyping, η οποία δέχεται ως είσοδο μη δομημένο κείμενο από τον χρήστη και το επεξεργάζεται με σκοπό να εξάγει την κρυμμένη γεωγραφική πληροφορία. Οι γεωγραφικοί όροι που εντοπίζονται αντιστοιχίζονται σε γεωγραφικές συντεταγμένες και απεικονίζονται σε χάρτη. Η εφαρμογή λειτουργεί σε πραγματικό χρόνο και εξασφαλίζει την αντιστοιχία μεταξύ γεωγραφικών όρων στον κείμενο και σημείων στον χάρτη.

Το Geotyping στοχεύει στην καλύτερη δυνατή αξιοποίηση του περιεχομένου που προέρχεται από απλούς χρήστες του διαδικτύου (user contributed content) μέσα από το πρίσμα της γεωγραφικής πληροφορίας που αυτό εμπεριέχει. Οι χρήστες καθημερινά δημοσιεύουν περιεχόμενο σε ιστολόγια, σελίδες κοινωνικής δικτύωσης και προσωπικές ιστοσελίδες. Το περιεχόμενο αυτό, αποτελεί μια μεγάλη βάση δεδομένων που σχετίζεται άμεσα με τον πραγματικό κόσμο. Η εφαρμογή συμβάλει στην ευρύτερη προσπάθεια εξόρυξης πληροφορίας από την γνώση που μοιράζονται οι άνθρωποι μεταξύ τους, με απώτερο στόχο την πιστή ψηφιοποίηση του πραγματικού κόσμου.

Εκτός από το Geotyping, σχεδιάσαμε και αναπτύξαμε ένα πλήρες σύστημα διαχείρισης κειμένου, το Geotyping Text Management System, το οποίο βασίζεται στο στην εφαρμογή Geotyping. Μέσω του συστήματος αυτού, ο εξουσιοδοτημένος χρήστης μπορεί να δημοσιεύει άρθρα τα οποία υπόκεινται σε επεξεργασία για τον εντοπισμό και την εξόρυξη γεωγραφικής πληροφορίας. Ο επισκέπτης προβάλλει ένα άρθρο, διαβάζοντας το κείμενο του και βλέποντας τους γεωγραφικούς όρους που περιείχε στον χάρτη.

1.2.1 Συνεισφορά

Για την ανάπτυξη της συγκεκριμένης εφαρμογής ακολουθήσαμε τα παρακάτω βήματα:

- Μελέτη συστημάτων και υπηρεσιών (Google Maps, CKEditor, Geoparser)
- Ανάλυση των απαιτήσεων της εφαρμογής
- Σχεδιασμός και υλοποίηση της εφαρμογής Geotyping
- Ανάλυση και σχεδιασμός του συστήματος διαχείρισης κειμένου με παράλληλη ενσωμάτωση της εφαρμογής Geotyping
- Αξιολόγηση αποτελεσμάτων για τα δύο συστήματα

1.3 Οργάνωση κειμένου

Η διπλωματική εργασία αποτελείται από 9 κεφάλαια. Το Κεφάλαιο 2 περιλαμβάνει τεχνολογίες και εργασίες σχετικές με το αντικείμενο της διπλωματικής. Στο Κεφάλαιο 3 αναλύονται οι απαιτήσεις των συστημάτων και περιγράφονται οι λειτουργίες τους. Στα Κεφάλαια 4 και 5 περιγράφεται η αρχιτεκτονική, η οργάνωση και τα βασικά χαρακτηριστικά του κάθε συστήματος. Στο κεφάλαιο 6 αναλύονται οι λεπτομέρειες υλοποίησης και τα προγραμματιστικά εργαλεία που χρησιμοποιήθηκαν. Το Κεφάλαιο 7 αποτελεί τον οδηγό χρήσης και εγκατάστασης των συστημάτων. Τέλος στο Κεφάλαιο 8, ανακεφαλαιώνουμε όλα τα θέματα με τα οποία ασχολήθηκε η παρούσα διπλωματική εργασία και στο Κεφάλαιο 9 παραθέτουμε την σχετική βιβλιογραφία που χρησιμοποιήθηκε για την συγγραφή του τόμου.

2. Σχετικές εργασίες

Το σύστημα Geotyping έρχεται να "δέσει" τις σύγχρονες υπηρεσίες γεωκωδικοποίησης και αναγνώρισης γεωγραφικών όρων σε ελεύθερο κείμενο, ώστε να συμβάλλει στην εξαγωγή και στην ταυτόχρονη απεικόνιση της γεωγραφικής πληροφορίας που εντοπίζεται. Στο παρόν κεφάλαιο παρουσιάζονται ορισμοί και εργασίες που σχετίζονται με το αντικείμενο της διπλωματικής εργασίας.

2.1 Ορισμοί

2.1.1 Geoparsing

Ως Geoparsing ορίζεται η λεκτική και γραμματική ανάλυση κειμένου με στόχο την ανακάλυψη κρυμμένης γεωγραφικής πληροφορίας. Η διαδικασία συμβάλλει στην κατανόηση του περιεχομένου του κειμένου και της γεωγραφικής πληροφορίας που αυτό περιέχει. Η γεωγραφική πληροφορία μπορεί να είναι είτε απόλυτη είτε σχετική. Για παράδειγμα η φράση "Είκοσι χιλιόμετρα δυτικά της Αθήνας" περιέχει απόλυτη γεωγραφική πληροφορία, την λέξη "Αθήνα" αλλά και σχετική πληροφορία, τον γεωγραφικό προσδιορισμό "Είκοσι χιλιόμετρα δυτικά" που σχετίζεται με την λέξη "Αθήνα". Η διαδικασία της λεκτικής και γραμματικής ανάλυσης μπορεί να επεκταθεί και σε ηχητικά αποσπάσματα λειτουργώντας με την ίδια ακριβώς μεθοδολογία. Το Geoparsing εντάσσεται στο γενικότερο πεδίο της Γεωγραφικής Ανάκτησης Πληροφοριών (Geographic Information Retrieval).

Σε συνδυασμό με μηχανισμούς γεωκωδικοποίησης, δημιουργούνται συστήματα τα οποία δέχονται ως είσοδο μη δομημένο κείμενο και εξάγουν τις γεωγραφικές συνταγμένες των γεωγραφικών όρων που εντοπίστηκαν στο κείμενο. Οι συντεταγμένες που προκύπτουν μπορούν να χαρτογραφηθούν σε Γεωγραφικά Συστήματα Πληροφοριών (Geographical Information Systems - GIS).

2.1.2 Γεωκωδικοποίηση

Η γεωκωδικοποίηση (geocoding) είναι η μετατροπή πληροφορίας σε συντεταγμένες σε κάποιο γεωγραφικό σύστημα αναφοράς. Μια απλή μέθοδος γεωκωδικοποίησης είναι η παρεμβολή διευθύνσεων (address interpolation) η οποία χρησιμοποιεί δεδομένα από γεωγραφικά συστήματα πληροφοριών στα οποία το οδικό δίκτυο αλλά και σημεία ενδιαφέροντος (Points Of Interest) για παράδειγμα μουσεία, αξιοθέατα, κυβερνητικά κτήρια, επιχειρήσεις και άλλα, είναι αντιστοιχισμένα σε γεωγραφικό χώρο συντεταγμένων. Το κάθε τμήμα δρόμου περιέχει το εύρος διευθύνσεων που του αναλογεί (για παράδειγμα τους αριθμούς των κτηρίων). Ο geocoder λαμβάνοντας ως είσοδο μια διεύθυνση, εντοπίζει το τμήμα του δρόμου στον οποίο ανήκει η διεύθυνση και παρεμβάλει την θέση της στο τμήμα αυτό.

Σήμερα, υπάρχουν διαθέσιμες αρκετές δωρεάν υπηρεσίες γεωκωδικοποίησης όπως η Google Maps, η Bing Maps, η OpenStreetMaps και άλλες. Η διαδικασία της γεωκωδικοποίησης μπορεί να αντιστραφεί. Η αντίστροφη γεωκωδικοποίηση δέχεται ως είσοδο τις συντεταγμένες ενός σημείου στο χάρτη και επιστρέφει την διεύθυνση του σημείου αυτού.

2.1.3 Geotagging

Ο όρος Geotagging αναφέρεται στην διαδικασία προσθήκης γεωγραφικών μεταδεδομένων ταυτοποίησης σε διάφορα μέσα, όπως φωτογραφίες, βίντεο, ιστοσελίδες, μηνύματα SMS, ή RSS feeds. Τα δεδομένα αυτά αποτελούν μια μορφή γεωχωρικών μεταδεδομένων και συνήθως αποτελούνται από γεωγραφικές συντεταγμένες, πληροφορίες για το υψόμετρο, την απόσταση, την ακρίβεια των δεδομένων, και τοπωνύμια.

Το Geotagging μπορεί να βοηθήσει τους χρήστες να βρουν μια μεγάλη ποικιλία δεδομένων για την περιοχή στην οποία βρίσκεται (Location aware services). Για παράδειγμα, ο χρήστης μπορεί να βρει εικόνες που λαμβάνονται κοντά σε μια δεδομένη τοποθεσία και να έχει πρόσβαση σε ειδήσεις, ιστοσελίδες και άρθρα που συνδέονται με μια περιοχή. Η σύνδεση μιας φωτογραφίας και των γεωγραφικών μεταδεδομένων της γίνεται είτε κατά την φάση της αποθήκευσης, είτε μέσω φωτογραφικών μηχανών και έξυπνων κινητών τηλεφώνων με ενσωματωμένο GPS δέκτη.

2.1.4 Geoblogging

Ο όρος Geoblogging (γεω-ιστολόγιο) χρησιμοποιείται για να περιγράψει μια εκδοχή ιστολογίου, που επιτρέπει την προσθήκη περιεχομένου το οποίο έχει γεωγραφική διάσταση και εμπεριέχει γεωγραφική πληροφορία. Η σύνδεση των ψηφιακών αντικειμένων του διαδικτύου (φωτογραφίες, video, ιστοσελίδες ή ιστολόγια) με την γεωγραφική θέση του στον παγκόσμιο χάρτη είναι πλέον ευρέως διαδεδομένη όπως αναφέρθηκε παραπάνω (Geotagging).

2.1.5 Ajax

Η τεχνολογία Ajax (Asynchronous JavaScript) είναι ένα σύνολο αλληλένδετων μεθόδων ανάπτυξης ιστοσελίδων που χρησιμοποιούνται στην πλευρά του πελάτη (client side) για να δημιουργήσουν διαδραστικές web εφαρμογές. Μέσω της τεχνολογίας αυτής οι εφαρμογές μπορούν να στείλουν και να λάβουν δεδομένα από έναν διακομιστή μέσω ασύγχρονων Http αιτημάτων χωρίς να απαιτείται η πλήρης επαναφόρτωση της σελίδας. Για την μεταφορά δεδομένων χρησιμοποιούνται οι τεχνολογίες XML και JSON.

Στον ορισμό δώσαμε τον όρο σύνολο μεθόδων ανάπτυξης καθώς η τεχνολογία Ajax χρησιμοποιεί ένα συνδυασμό Html και Css για την σήμανση, την απεικόνιση και την διαχείριση των πληροφοριών. Ακόμη, το DOM (Document Object Model) σε μια ιστοσελίδα, που είναι προσβάσιμο μέσω της JavaScript, δίνει την δυνατότητα δυναμικής εμφάνισης δεδομένων και επιτρέπει στον χρήστη να αλληλεπιδρά με τις πληροφορίες που παρουσιάζονται.

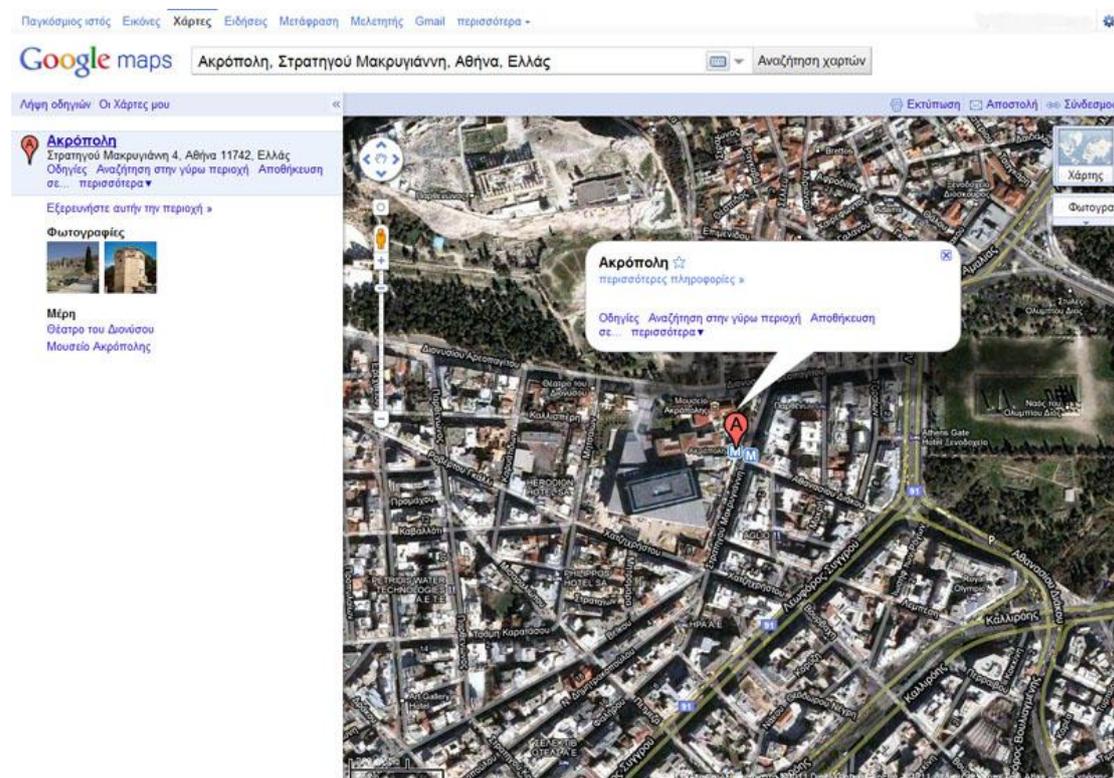
2.2 Σχετικές Εργασίες

2.2.1 Google Maps

Το Google Maps είναι μια ελεύθερη υπηρεσία διαδικτυακής χαρτογράφησης (web mapping) που παρέχεται από την Google και τροφοδοτεί με χάρτες υψηλής ανάλυσης πολλές ιστοσελίδες τρίτων μέσω του Google Maps API. Η υπηρεσία προσφέρει οδικό δίκτυο, δορυφορικές φωτογραφίες, λειτουργία για σχεδιασμό διαδρομών, εντοπισμό επιχειρήσεων καθώς και τρισδιάστατη εικονική περιήγηση σε πολλές περιοχές.

Με τους χάρτες Google είναι δυνατή η πλοήγηση του χρήστη στο χάρτη με το σύρσιμο του ποντικιού, αλλά και η αναζήτηση κάποιας γεωγραφικής θέσης βάσει κάποιας διεύθυνσης, ενός ταχυδρομικού κώδικα ή κάποιας λεκτικά ορισμένης περιοχής. Ακόμη, παρέχουν οδηγίες πλοήγησης, παραθέτοντας στο χρήστη μια λίστα με μεμονωμένες κινήσεις για το πώς να φτάσει στον προορισμό του, μαζί με μια εκτίμηση του χρόνου που απαιτείται για αυτή την απόσταση.

Το Google Maps API που παρέχεται μαζί με την υπηρεσία, δίνει την δυνατότητα πλήρους παραμετροποίησης του χάρτη και του Geocoder που θα ενσωματωθεί στην εκάστοτε εφαρμογή. Ένα στιγμιότυπο της εφαρμογής δίνεται στην εικόνα 2.1.



2-1 Η υπηρεσία Google Maps

2.2.2 *Yahoo Placemaker*

Το Yahoo Placemaker είναι μια ελεύθερη υπηρεσία geoparsing που στοχεύει στην δημιουργία εφαρμογών με επίγνωση θέσης. Η λειτουργία του συνοψίζεται στον εντοπισμό των γεωγραφικών όρων που βρίσκονται σε αδόμητο περιεχόμενο, RSS feeds, ιστοσελίδες, ειδήσεις και στην εξόρυξη των γεωγραφικών μεταδεδομένων που θα χρησιμοποιηθούν για την γεωγραφική επισήμανση και την κατηγοριοποίηση του περιεχομένου αυτού. Η υπηρεσία ουσιαστικά εμπλουτίζει ένα σημαντικό ποσοστό περιεχομένου του διαδικτύου το οποίο περιέχει κρυμμένη γεωγραφική πληροφορία που δεν έχει εντοπιστεί. Το σύστημα περιλαμβάνει Geoparser και Geocoder.

2.2.3 *Geocrowd*

Το Geocrowd είναι μια υπηρεσία γεω-ιστολογίου (geoblogging) που χρησιμοποιείται για την χωροχρονική αφήγηση ενός ταξιδιού όπως είναι μια απογευματινή βόλτα, μια πεζοπορία ή μια διαδρομή με το αυτοκίνητο. Μέσω της εφαρμογής, ο χρήστης μπορεί να αφηγηθεί την ιστορία του ταξιδιού με βάση το γεωγραφικό περιεχόμενο που συλλέχθηκε κατά την διάρκεια του. Στόχος της εφαρμογής είναι η συγκέντρωση γεωγραφικών δεδομένων τα οποία βασίζονται σε συγκεκριμένες εμπειρίες και βοηθούν στην πιστή ψηφιοποίηση του πραγματικού κόσμου.

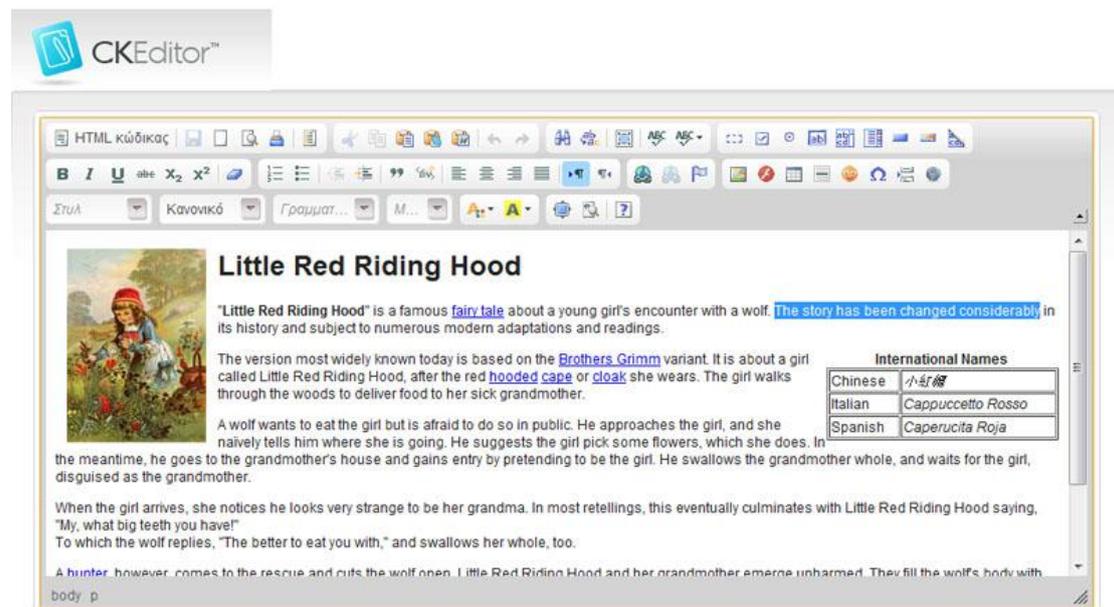
2.2.4 *Γεωκωδικοποίηση ιστοσελίδων*

Η αναζήτηση και η κατηγοριοποίηση των ιστοσελίδων μπορεί να πραγματοποιηθεί μέσω του περιεχόμενου της ιστοσελίδας, με την χρήση λέξεων κλειδιών αλλά και με τεχνικές που λαμβάνουν υπ όψιν την γεωγραφική ταυτότητα της σελίδας. Ουσιαστικά πρόκειται για την γεωκωδικοποίηση της ιστοσελίδας και την αντιστοίχισής της με συγκεκριμένους γεωγραφικούς όρους. Τα τελευταία χρόνια έχουν αναπτυχθεί ημιαυτόματες τεχνικές γεωκωδικοποίησης ιστοσελίδων που προσπαθούν να εντοπίσουν αριθμούς τηλεφώνων και διευθύνσεις με σκοπό την εξαγωγή του γεωγραφικών συμπερασμάτων. Λόγω της ασάφειας της γλώσσας, απαιτείται και η δημιουργία εργαλείων για την χειροκίνητη γεωκωδικοποίηση ιστοσελίδων.

2.2.5 CKEditor

Ο CKEditor είναι ένα online πρόγραμμα επεξεργασίας κειμένου (Html editor) γραμμένο σε JavaScript, που ενσωματώνεται σε ιστοσελίδες όπως είναι τα συστήματα διαχείρισης περιεχομένου και τα blogs. Είναι ένα πρόγραμμα επεξεργασίας WYSIWYG, πράγμα που σημαίνει ότι το κείμενο υπό επεξεργασία πρέπει να φαίνεται όσο το δυνατόν όμοιο με το κείμενο που δημοσιεύει ο χρήστης. Οι HTML editors επεκτείνουν την online επεξεργασία κειμένου με πολλά από τα χαρακτηριστικά που υπάρχουν σε stand alone εφαρμογές επεξεργασίας κειμένου όπως το Microsoft Word και το Open Office.

Ο CKEditor είναι πρόγραμμα ανοιχτού κώδικα ενώ παρέχει παράλληλα εμπορικές άδειες. Συνεπώς μπορεί να χρησιμοποιηθεί σε οποιαδήποτε εφαρμογή και να παρέχει εύκολες και δυναμικές λύσεις στους χρήστες και στους προγραμματιστές. Επιπρόσθετα, συνοδεύεται από ένα πλούσιο API για την διαχείριση του επεξεργαστή κειμένου και των δεδομένων. Η προσθήκη νέων λειτουργιών και επεκτάσεων υποστηρίζεται πλήρως, καθώς ο προγραμματιστής μπορεί να υλοποιήσει τις λειτουργίες που επιθυμεί και να τις εντάξει στο γραφικό περιβάλλον του CKEditor έχοντας στην διάθεση του μια μεγάλη γκάμα εργαλείων και λειτουργιών μέσω του API.



2-2 Γραφικό περιβάλλον του CKEditor

2.2.6 jQuery

Η jQuery είναι η δημοφιλέστερη cross-browser JavaScript βιβλιοθήκη. Σχεδιάστηκε για να απλοποιεί τον κώδικα που απαιτείται στην πλευρά του πελάτη. Η βιβλιοθήκη είναι ανοιχτού κώδικα και στοχεύει στο να διευκολύνει τον προγραμματιστή στην διαχείριση των στοιχείων του DOM, στην δημιουργία γραφικών και animations, στην διαχείριση γεγονότων και στην αποστολή Ajax αιτημάτων προς τον διακομιστή.

2.2.7 Σύστημα διαχείρισης κειμένου

Τα συστήματα διαχείρισης κειμένου αποτελούν υποσύνολο των συστημάτων διαχείρισης περιεχομένου. Ένα σύστημα διαχείρισης περιεχομένου ορίζεται ως ένα σύνολο διαδικασιών που χρησιμοποιούνται για τη διαχείριση της ροής εργασίας σε ένα συνεργατικό περιβάλλον. Οι διαδικασίες αυτές μπορεί να είναι είτε χειροκίνητες είτε να βασίζονται σε υπολογιστή. Τα χαρακτηριστικά του συστήματος είναι τα εξής:

- Επιτρέπει σε έναν μεγάλο αριθμό χρηστών να αποθηκεύσουν και να μοιραστούν δεδομένα κειμένου
- Έλεγχος πρόσβασης στα δεδομένα, βασισμένος στους ρόλους χρηστών. Το σύστημα ορίζει ομάδες χρηστών οι οποίες μπορούν να προβάλλουν , να επεξεργαστούν αλλά και να δημοσιεύσουν δεδομένα.
- Παρέχει εργαλεία επικοινωνίας μεταξύ των χρηστών.
- Παρέχει δυνατότητα ταξινόμησης και κατηγοριοποίησης του περιεχομένου.

Παραδείγματα δημοφιλών συστημάτων διαχείρισης περιεχομένου ανοιχτού κώδικα αποτελούν οι πλατφόρμες Joomla, Drupal και Wordpress. Τα συστήματα αυτά διαχειρίζονται πολλαπλούς τύπους πολυμέσων, επομένως ένα σύστημα διαχείρισης κειμένου διαχειρίζεται γραπτό κείμενο.

3. *Ανάλυση Απαιτήσεων Συστήματος*

Το παρόν κεφάλαιο είναι αφιερωμένο στην περιγραφή της αρχιτεκτονικής και στην ανάλυση των απαιτήσεων της εφαρμογής Geotyping και του συστήματος διαχείρισης κειμένου Geotyping Text Management System. Αρχικά δίνεται η λεπτομερής περιγραφή των απαιτήσεων των εφαρμογών. Εν συνεχεία, τα συστήματα αναλύονται σε επιμέρους υποσυστήματα με διακριτό ρόλο. Τέλος, παρουσιάζονται οι βασικές λειτουργίες των εφαρμογών, σενάρια χρήσης καθώς και το διάγραμμα οντοτήτων συσχετίσεων για το σύστημα διαχείρισης κειμένου.

3.1 *Αρχιτεκτονική*

Το βασικό αντικείμενο της παρούσας διπλωματικής είναι η σχεδίαση και η ανάπτυξη της εφαρμογής Geotyping με στόχο την γεωκωδικοποίηση ελεύθερου κειμένου σε πραγματικό χρόνο. Το σύστημα διαχείρισης κειμένου Geotyping Text Management System σχεδιάστηκε ώστε να δώσει ένα παράδειγμα πρότυπου τρόπου ενσωμάτωσης και αξιοποίησης των δυνατοτήτων της εφαρμογής Geotyping σε ένα ολοκληρωμένο περιβάλλον.

3.1.1 *Geotyping*

Το Geotyping είναι μια Ajax client-server εφαρμογή η οποία δέχεται ως είσοδο μη δομημένο κείμενο που πληκτρολογεί ο χρήστης και παρέχει ως έξοδο τις συντεταγμένες των γεωγραφικών όρων που εντοπίστηκαν στο κείμενο. Οι συντεταγμένες απεικονίζονται με δείκτες στον χάρτη.

Επιγραμματικά, η λειτουργία της εφαρμογής ξεκινά με την πληκτρολόγηση κειμένου στον HTML editor (CKEditor) από τον χρήστη. Στην συνέχεια, η εφαρμογή αναλαμβάνει να επεξεργαστεί την προηγούμενη παράγραφο του κειμένου και να την στέλνει στον Geoparser. Η επικοινωνία γίνεται μέσω ενός τοπικού proxy server που αναλαμβάνει να προωθεί τα αιτήματα στον εξωτερικό Geoparser της εφαρμογής. Ο Geoparser, αφού επεξεργαστεί το κείμενο εισόδου μέσω τεχνικών επεξεργασίας φυσικής γλώσσας, επιστρέφει στην εφαρμογή

τους πιθανούς γεωγραφικούς όρους που εντοπίστηκαν. Το Geotyping με την σειρά του αντιστοιχίζει και απεικονίζει τους γεωγραφικούς όρους με συντεταγμένες στον χάρτη μέσω της υπηρεσίας γεωκωδικοποίησης του Google Maps.

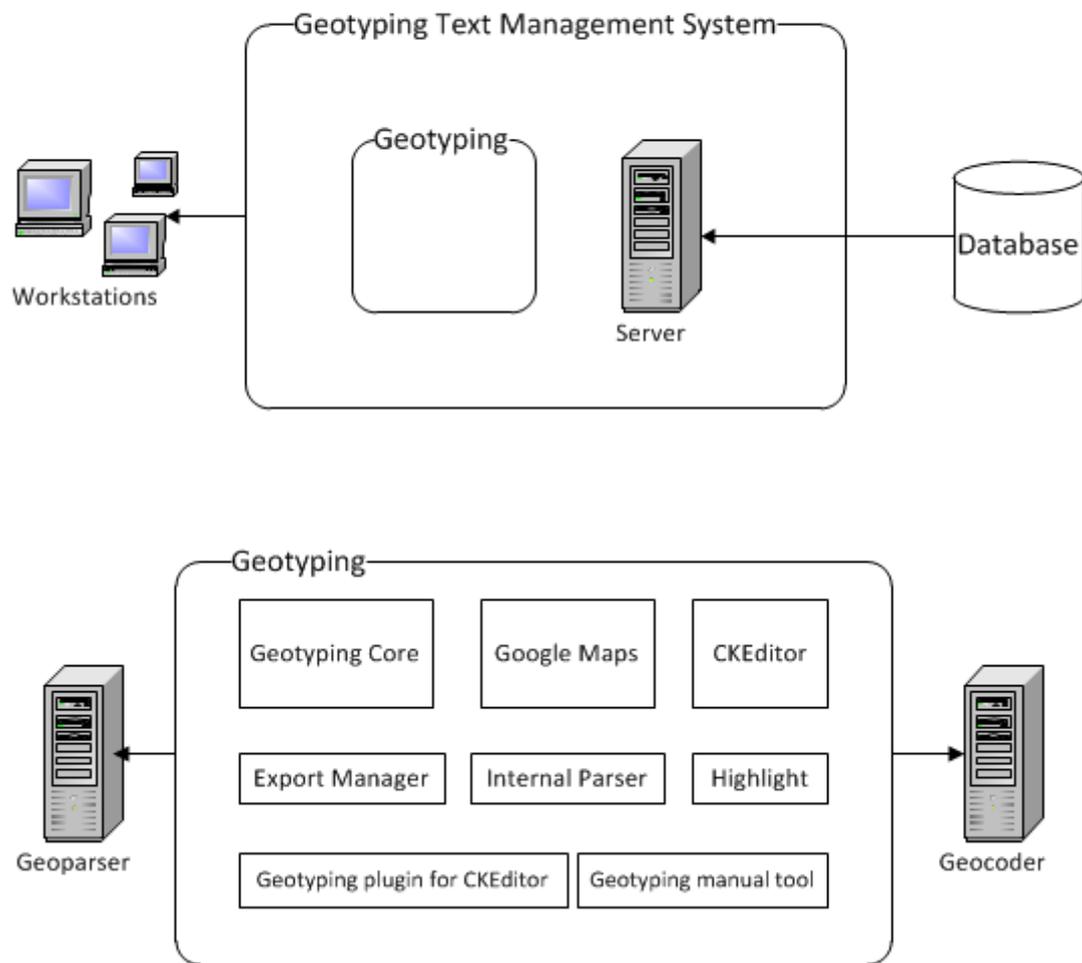
Η εφαρμογή είναι πλήρως διαδραστική. Η πληκτρολόγηση και η γεωγραφική απεικόνιση γίνονται ταυτόχρονα. Επιπλέον, ο χρήστης έχει την δυνατότητα να σύρει μια οποιαδήποτε φράση στον χάρτη και να την τοποθετήσει στο σημείο που επιθυμεί, αντιστοιχίζοντας το κείμενο της φράσης με το συγκεκριμένο σημείο.

Η βασικότερη απαίτηση της εφαρμογής είναι η αντιστοιχία που πρέπει να υπάρχει μεταξύ των σημείων στον χάρτη και των αντίστοιχων γεωγραφικών όρων στο κείμενο. Η ταχύτητα επεξεργασίας της προηγούμενης παραγράφου που εισήγαγε ο χρήστης, ώστε η εφαρμογή να είναι πραγματικού χρόνου, η μαζική γεωκωδικοποίηση φράσεων και λέξεων καθώς επίσης η δημιουργία ενός εύχρηστου γραφικού περιβάλλοντος αποτέλεσαν εξίσου σημαντικές προκλήσεις κατά την διαδικασία της ανάπτυξης.

Αναλύοντας την εφαρμογή σε επιμέρους υποσυστήματα, όπως φαίνεται στο διάγραμμα του σχήματος 4.3, γίνεται εύκολα αντιληπτό ότι το Geotyping αποτελείται από τρία βασικά υποσυστήματα και αρκετά εργαλεία τα οποία ενισχύουν τον ρόλο των βασικών υποσυστημάτων.

Συνοψίζοντας, η εφαρμογή καταφέρνει να εξάγει σε πραγματικό χρόνο, την κρυμμένη γεωγραφική πληροφορία του κειμένου του χρήστη καθώς αυτό πληκτρολογείται. Στο σύστημα διαχείρισης κειμένου, που αναλύεται σε επόμενη ενότητα, αποθηκεύουμε το κείμενο του χρήστη μαζί με τις συντεταγμένες των γεωγραφικών όρων που εντοπίστηκαν σε μια βάση δεδομένων δίνοντας ένα παράδειγμα αξιοποίησης της εφαρμογής.

Διάγραμμα υποσυστημάτων



3-1 Διάγραμμα υποσυστημάτων

3.1.2 Περιγραφή Λειτουργιών

3.1.2.1 Geotyping Core

Πρόκειται για το βασικότερο υποσύστημα της εφαρμογής. Οι ενέργειες που περιλαμβάνει χωρίζονται σε τρεις κατηγορίες:

- **Αρχειοποίηση και ρύθμιση παραμέτρων**

Το υποσύστημα αναλαμβάνει να αρχειοποιήσει όλα τα υπόλοιπα υποσυστήματα κατά την φόρτωση της εφαρμογής. Επεξεργάζεται τις ρυθμίσεις που ορίζει ο χρήστης και θέτει κατάλληλα τις παραμέτρους λειτουργίας της εφαρμογής.

- **Διαχείριση της ροής των δεδομένων**

Το υποσύστημα στέλνει την προηγούμενη παράγραφο του κειμένου στον Geoparser, αναμένει την απάντηση του και στην συνέχεια προωθεί τους πιθανούς γεωγραφικούς όρους που εντοπίστηκαν στον Geocoder. Επιπρόσθετα, είναι ικανό να διαχειριστεί τα σφάλματα που πιθανών προκύψουν κατά την διάρκεια της επικοινωνίας των επιμέρους συστημάτων.

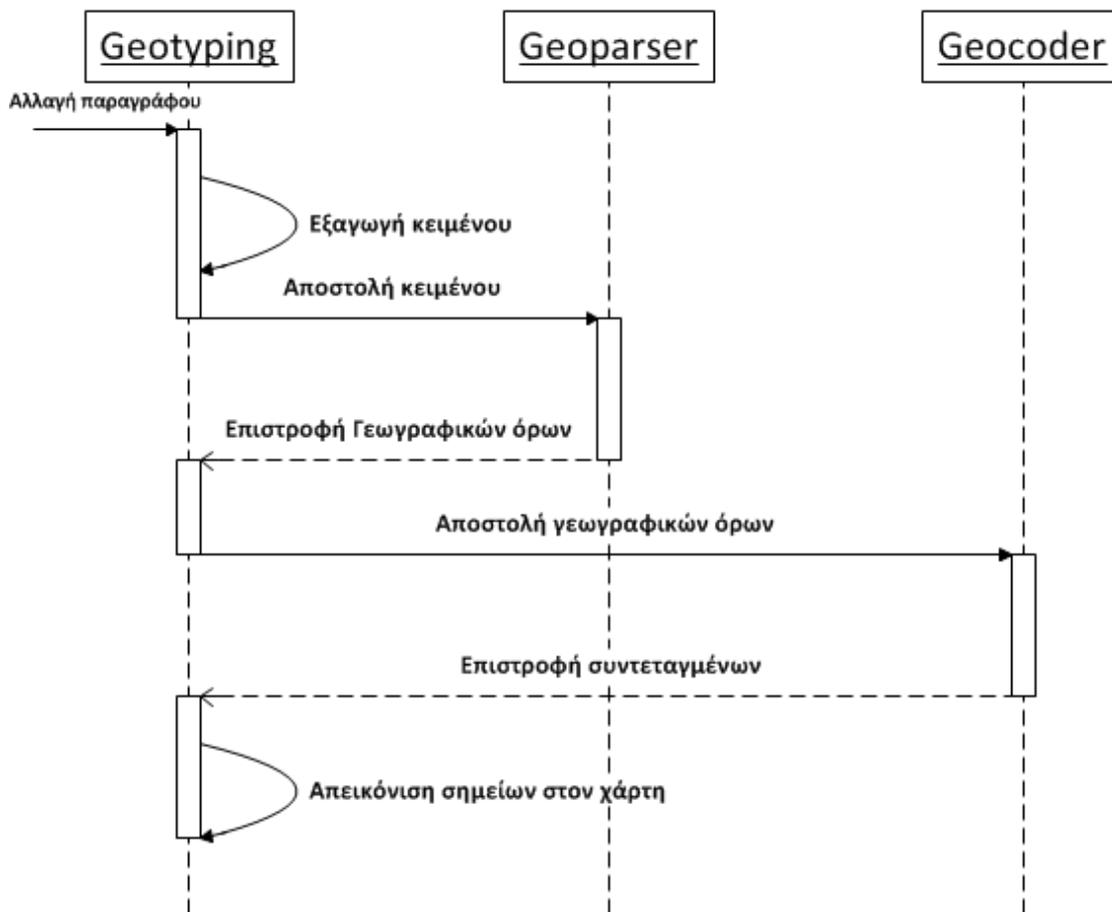
Η αποθήκευση των γεωγραφικών όρων με τις αντίστοιχες συντεταγμένες τους γίνεται στην δομή Geotyping Hash. Κατά την ροή των δεδομένων, χρησιμοποιούνται προσωρινά μέσα αποθήκευσης (buffers) ενδιάμεσων αποτελεσμάτων που στοχεύουν στην μείωση των αιτημάτων προς τον Geoparser και τον Geocoder. Το σημείο αυτό απαιτεί την μεγαλύτερη προσοχή στην ανάπτυξη της εφαρμογής μιας και αποτελεί το σημαντικότερο σημείο συμφόρησης (bottleneck).

Στο ακολουθιακό διάγραμμα που δίνεται στο σχήμα 4 απεικονίζεται η ροή των δεδομένων ανάμεσα στα υποσυστήματα. Ένα σημείο στο οποίο θα σταθούμε σε επόμενο κεφάλαιο είναι η έννοια της προηγούμενης παραγράφου στο κείμενο που πληκτρολογεί ο χρήστης στον CKEditor.

- **Αλληλεπίδραση με τον CKEditor**

Το υποσύστημα επεκτείνει τις βασικές λειτουργίες του CKEditor, εξάγει κείμενο από αυτόν και αποθηκεύει τις ρυθμίσεις του χρήστη μέσω του Geotyping Plug-in For CKEditor.

Ακολουθιακό διάγραμμα



3-2 Ακολουθιακό διάγραμμα Geotyping

3.1.2.2 Google Maps

Το υποσύστημα περιλαμβάνει τον χάρτη και τον Geocoder. Είναι υπεύθυνο για την γεωκωδικοποίηση και την παρουσίαση των γεωγραφικών συντεταγμένων στον χάρτη. Η απεικόνιση των όρων γίνεται μέσω δεικτών (markers). Ο χρήστης έχει την δυνατότητα αλληλεπίδρασης με τους δείκτες καθώς μπορεί να διαγράψει ένα όρο από τον χάρτη ή να αλλάξει την θέση του στην περίπτωση που ο Geocoder επέστρεψε διαφορετική θέση από την αναμενόμενη. Κάνοντας κλικ πάνω στον δείκτη εμφανίζεται ένα πλαίσιο (info window) που περιέχει την διεύθυνση του συγκεκριμένου σημείου και τα κουμπιά επεξεργασίας και διαγραφής του δείκτη. Η δημιουργία και η απεικόνιση των πλαισίων και των δεικτών αποτελεί ευθύνη του υποσυστήματος Google Maps.

Επιπρόσθετα, το υποσύστημα αναλαμβάνει την γεωκωδικοποίηση και την αντίστροφη γεωκωδικοποίηση γεωγραφικών όρων. Κατά την λειτουργία της

γεωκωδικοποίησης, το σύστημα δέχεται σαν είσοδο μια φράση - γεωγραφικό όρο, και επιστρέφει την θέση του στον χάρτη σε ένα προκαθορισμένο σύστημα συντεταγμένων. Κατά την αντίστροφη γεωκωδικοποίηση, το υποσύστημα δέχεται ως είσοδο τις συντεταγμένες ενός σημείου στον χάρτη και επιστρέφει την διεύθυνση του. Η λειτουργία της αντίστροφης γεωκωδικοποίησης είναι απαραίτητη για την διόρθωση ενός σημείου στον χάρτη. Ο χρήστης επιλέγει να διορθώσει τον δείκτη σέρνοντας με το ποντίκι το εικονίδιο του σε μια νέα θέση η οποία μεταφράζεται σε νέες συντεταγμένες.

3.1.2.3 CKEditor

Ο CKEditor αποτελεί έναν ευρέως χρησιμοποιούμενο HTML editor καθώς δίνει πολλές δυνατότητες στην online επεξεργασία κειμένου μέσω μιας μεγάλης γκάμας εργαλείων που διαθέτει. Λόγω των χαρακτηριστικών του επιλέχθηκε να ενσωματωθεί στην εφαρμογή. Περισσότερες πληροφορίες σχετικά με τον CKEditor μπορείτε να βρείτε στην ιστοσελίδα της εφαρμογής.

3.1.2.4 Manual Tool

Το υποσύστημα προσφέρει την δυνατότητα χειροκίνητης αντιστοίχισης φράσεων με συντεταγμένες στον χάρτη. Ο χρήστης επιλέγει στο κείμενο την φράση που επιθυμεί ενεργοποιώντας την χειροκίνητη λειτουργία. Στην συνέχεια τοποθετεί το ποντίκι του στο επιθυμητό σημείο στον χάρτη ολοκληρώνοντας την αντιστοίχιση.

3.1.2.5 Geoparser

Η ανάλυση του εξωτερικού Geoparser ξεπερνάει τα πλαίσια της παρούσας διπλωματικής εργασίας. Περιγράφοντας επιγραμματικά την λειτουργία του, πρόκειται για ένα σύστημα φυσικής επεξεργασίας γλώσσας (Natural Language Processing) το οποίο αναλύει ελεύθερο κείμενο χωρίζοντας το σε προτάσεις, λέξεις και μέρη του λόγου χρησιμοποιώντας το General Architecture for Text Engineering Framework (GATE). Λαμβάνοντας τα αποτελέσματα της παραπάνω επεξεργασίας, πραγματοποιεί σημασιολογική ανάλυση, συνδυάζοντας γλωσσική πληροφορία με οντολογίες και γλωσσικούς κανόνες ώστε να μπορέσει να εξάγει γεωγραφική πληροφορία, μέσω του συστήματος CAFETIERE. Μετά την ολοκλήρωση της επεξεργασίας, το υποσύστημα επιστρέφει τους υποψήφιους γεωγραφικούς όρους που εντοπίστηκαν. Περισσότερες πληροφορίες σχετικά με τον Geoparser μπορείτε να βρείτε στην επισυναπτόμενη βιβλιογραφία.

3.1.2.6 Εσωτερικός parser

Πρόκειται για μια απλοϊκή υλοποίηση ενός εσωτερικού γρήγορου parser ο οποίος αναγνωρίζει μεμονωμένες λέξεις και ακολουθίες λέξεων με κεφαλαίο γράμμα σε ελληνικό και αγγλικό κείμενο. Χρησιμοποιήθηκε κυρίως στο στάδιο της ανάπτυξης της εφαρμογής και παραμένει διαθέσιμος στον τελικό χρήστη μέσα από της επιλογές του Geotyping Plug-in for CKEditor. Όπως αναφέρθηκε, πρόκειται για μια απλοϊκή υλοποίηση ενός αναλυτική κειμένου που πλεονεκτεί στον χρόνο απόκρισης σε σύγκριση με τον εξωτερικό parser.

3.1.2.7 Geotyping Hash

Πρόκειται για μια δομή δεδομένων, έναν πίνακα κατακερματισμού (hash table), στον οποίο αποθηκεύονται οι γεωγραφικοί όροι του κειμένου με τις συντεταγμένους τους. Τα δεδομένα βρίσκονται αποθηκευμένα όσο εκτελείται η εφαρμογή και μπορούν να εξαχθούν μέσω του υποσυστήματος Export Manager. Η δομή προσφέρει την δυνατότητα γραφικής απεικόνισης της σε μορφή λίστας ώστε ο χρήστης να έχει μια εποπτική εικόνα για τους γεωγραφικούς όρους που εντοπίστηκαν στο κείμενο.

3.1.2.8 Highlighter

Το υποσύστημα αναλαμβάνει να επισημάνει στο κείμενο τις φράσεις-γεωγραφικούς όρους που αντιστοιχούν σε σημεία στον χάρτη. Η επισήμανση γίνεται με κίτρινο φόντο. Κατά την επεξεργασία προηγούμενων παραγράφων όπου ο χρήστης πραγματοποιεί διορθώσεις στο κείμενο, το υποσύστημα αφαιρεί την επισήμανση και την επαναφέρει μετά την ολοκλήρωση της διαδικασίας όπου απαιτείται.

3.1.2.9 Proxy Server

Το υποσύστημα λειτουργεί εκ μέρους του εξωτερικού Geoparser, δεχόμενο αιτήσεις τις οποίες και προωθεί στον Geoparser. Η απαίτηση χρήσης ενός proxy server στην εφαρμογή θα αναλυθεί σε επόμενο κεφάλαιο.

3.1.2.10 Export Manager

Το υποσύστημα αναλαμβάνει να εξάγει όλα τα δεδομένα από την δομή Geotyping Hash σε μορφή XML ή JSON. Τα δεδομένα αυτά μπορούν να αποθηκευτούν σε αρχείο στον δίσκο ή σε βάση δεδομένων και αφού συσχετιστούν με το κείμενο από το οποίο εξήχθησαν να απεικονιστούν σε χάρτη. Η λειτουργία αυτή εκτελείται κατά την αποθήκευση μιας δημοσίευσης από το σύστημα διαχείρισης κειμένου.

3.1.2.11 Geotyping Plug-in for CKEditor

Πρόκειται για ένα ενσωματωμένο γραφικό περιβάλλον μέσω του οποίου μπορούμε να ρυθμίσουμε παραμέτρους του Geotyping και να προσθέσουμε μενού που εκτελούν συγκεκριμένες λειτουργίες. Αποτελείται από ένα αναδυόμενο παράθυρο στο οποίο μπορούμε να ρυθμίζουμε αρκετές παραμέτρους του Geotyping, και από ένα context menu στην περιοχή κειμένου του CKEditor, μέσω του οποίου μπορούμε να απεικονίσουμε χειροκίνητα μια φράση στον χάρτη, να διαγράψουμε έναν γεωγραφικό όρο, να προσθέσουμε έναν όρο στην μαύρη λίστα και να επανεκκινήσουμε την εφαρμογή.

3.1.3 Geotyping Text Management System

Το Geotyping Text Management System είναι ένα πλήρες σύστημα διαχείρισης κειμένου το οποίο σχεδιάστηκε ώστε να εκμεταλλεύεται πλήρως τις δυνατότητες της εφαρμογής Geotyping. Περιλαμβάνει σύστημα διαχείρισης χρηστών, προβολή και επεξεργασία άρθρων, ανάκτηση γεωγραφικών όρων με ταυτόχρονη απεικόνιση σε χάρτη και αναζήτηση άρθρων που βασίζεται σε μέτα ετικέτες.

Οι χρήστες - συντάκτες αφού εγγραφούν στο σύστημα, μπορούν να δημοσιεύσουν τα άρθρα τους μέσω του Geotyping και να διαχειριστούν τα στοιχεία του λογαριασμού τους. Κάθε άρθρο συσχετίζεται με τα γεωγραφικά δεδομένα που περιέχει και αποθηκεύεται στην βάση δεδομένων . Η προβολή των δημοσιευμένων άρθρων αποτελείται από το κείμενο και από έναν χάρτη στο οποίο έχουν σημειωθεί οι γεωγραφικοί όροι. Μέσω του συγκεκριμένου συστήματος δίνεται ένα παράδειγμα οπτικοποίησης την γεωγραφική πληροφορία που εξήχθη από το κείμενο. Ακολουθεί η αναλυτική περιγραφή των λειτουργιών του συστήματος διαχείρισης κειμένου.

3.1.4 Περιγραφή λειτουργιών

3.1.4.1 Διαχείριση χρηστών

Ο χρήστης για να μπορέσει να δημοσιεύσει τα άρθρα του, θα πρέπει να εγγραφεί στο σύστημα. Οι ανώνυμοι χρήστες μπορούν μόνο να προβάλουν τα δημοσιευμένα άρθρα και όχι να προσθέσουν νέα. Η εγγραφή γίνεται μέσω της φόρμας εγγραφής. Αφού ολοκληρωθεί η διαδικασία, ο εγγεγραμμένος χρήστης συνδέεται στον λογαριασμό του μέσω της φόρμας σύνδεσης έχοντας πλέον την δυνατότητα προσθήκης και διαχείρισης άρθρων. Το σύστημα προσφέρει ακόμη υπηρεσία ανάκτησης κωδικού, διαχείριση στοιχείων λογαριασμού και αντιμετώπιση των σφαλμάτων που πιθανώς προκύψουν. Τέλος εξασφαλίζει την

μοναδικότητα του ονόματος χρήστη και του email και την κρυπτογραφημένη αποθήκευση του κωδικού πρόσβασης.

3.1.4.2 Δημοσίευση και επεξεργασία άρθρων

Αποτελεί την βασικότερη λειτουργία της εφαρμογής. Οι συνδεδεμένοι χρήστες μπορούν να δημοσιεύσουν νέα άρθρα και να επεξεργαστούν παλαιότερες δημοσιεύσεις. Η δημοσίευση και η επεξεργασία γίνεται μέσω της εφαρμογής Geotyping. Σε κάθε άρθρο ορίζονται ετικέτες σχετικές με το περιεχόμενο του και αποτελούν τμήμα του μηχανισμού αναζήτησης άρθρων.

3.1.4.3 Αναζήτηση άρθρων

Το σύστημα παρέχει την δυνατότητα αναζήτησης άρθρων είτε βάση των μετα-ετικετών που συνδέονται με το κάθε άρθρο είτε με βάση του τίτλου κάθε άρθρου. Ο χρήστης εισάγει φράση που επιθυμεί και το σύστημα αναλαμβάνει να του επιστρέψει τα σχετικά αποτελέσματα.

3.1.4.4 Προβολή άρθρων

Η προβολή ενός άρθρου περιλαμβάνει το κείμενο και τον χάρτη. Στον χάρτη υπάρχουν οι δείκτες των γεωγραφικών όρων που έχουν επισημανθεί στο κείμενο.

3.2 Σενάρια χρήσης

Ακολουθώντας την μεθοδολογία των περιπτώσεων χρήσης για το σύστημα διαχείρισης κειμένου, αναγνωρίζουμε τους χρήστες-δράστες του συστήματος και καταγράφουμε για κάθε δράση τις απαιτούμενες ενέργειες του χρήστη.

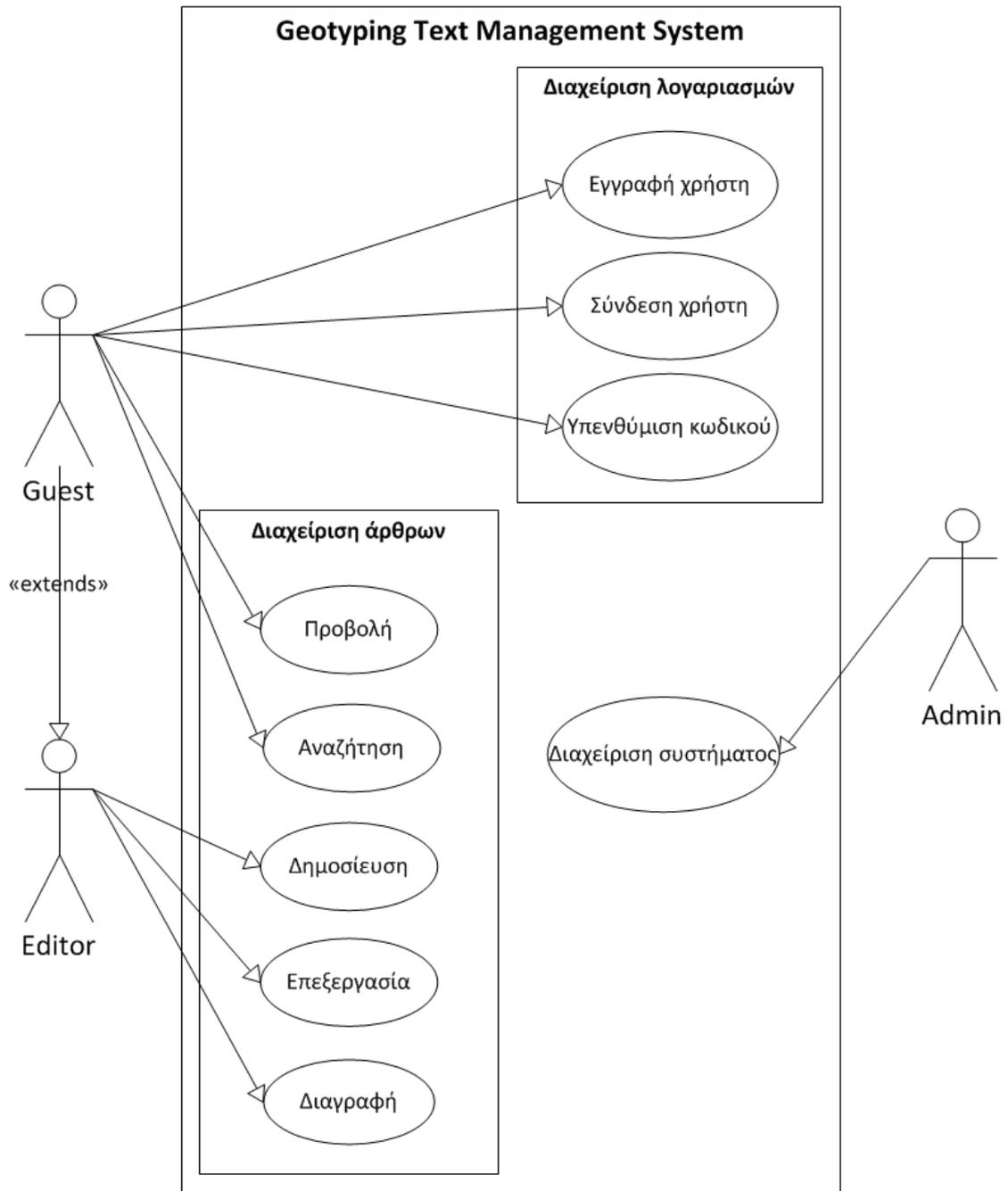
Στην εφαρμογή που αναπτύχθηκε για την διπλωματική εργασία, οι δράστες (actors) του συστήματος είναι τρεις, ο επισκέπτης, ο συντάκτης και ο διαχειριστής. Οι διακριτές δράσεις που δύναται να πραγματοποιήσει ο συντάκτης είναι οι ακόλουθες:

- Προβολή άρθρου
- Αναζήτηση άρθρου
- Είσοδος στο σύστημα
- Διαχείριση προφίλ λογαριασμού
- Δημοσίευση άρθρου
- Επεξεργασία δημοσιευμένου άρθρου
- Διαγραφή άρθρου

Ο επισκέπτης εκτελεί τις παρακάτω δράσεις:

- Προβολή άρθρου
- Αναζήτηση άρθρου
- Εγγραφή στο σύστημα

Παρατηρούμε ότι ο συντάκτης ουσιαστικά κληρονομεί και επεκτείνει τις δράσεις του επισκέπτη. Ακολουθεί το διάγραμμα χρήσης (Use Case Diagram) και η ανάλυση του βασικότερου σεναρίου χρήσης της εφαρμογής.



3-3 Διάγραμμα χρήσης

3.2.1 Περίπτωση χρήσης "Δημοσίευση νέου άρθρου"

Τίτλος περίπτωσης χρήσης:

Δημοσίευση νέου άρθρου

Σύντομη περιγραφή:

Ο χρήστης αφού συνδεθεί στο σύστημα, επιθυμεί να δημοσιεύσει ένα νέο άρθρο.

Ροή γεγονότων:**Βασική ροή:**

1. Ο χρήστης επιλέγει το σύνδεσμο σύνδεσης λογαριασμού.
2. Εισάγει τα στοιχεία του στην φόρμα σύνδεσης.
3. Μετά την επιτυχή σύνδεση, επιλέγει τον σύνδεσμο νέας δημοσίευσης στην αρχική σελίδα του συντάκτη.
4. Συμπληρώνει όλα τα πεδία που απαιτούνται για το νέο άρθρο, τίτλο, ετικέτες και εισάγει το περιεχόμενο του μέσω της εφαρμογής Geotyping η οποία παράλληλα, εξάγει σε πραγματικό χρόνο την κρυμμένη γεωγραφική πληροφορία.
5. Ολοκληρώνοντας της διαδικασία, επιλέγει δημοσίευση του άρθρου.

Εναλλακτικές ροές:

Σε περίπτωση αποτυχίας σύνδεσης, ο χρήστης μεταβαίνει στην σελίδα υπενθύμισης κωδικού.

Μη λειτουργικές απαιτήσεις:

Δεν υπάρχουν μη λειτουργικές απαιτήσεις για τη συγκεκριμένη περίπτωση χρήσης.

Κατάσταση εισόδου:

Ο χρήστης του συστήματος αναγνωρίζεται από το σύστημα ως συνδεδεμένος χρήστης με μοναδικό αναγνωριστικό.

Κατάσταση εξόδου:

Το νέο άρθρο αποθηκεύεται στην βάση και είναι διαθέσιμο για προβολή από τους επισκέπτες του συστήματος.

3.3 Μοντέλο Οντοτήτων Συσχετίσεων

Το σύστημα διαχείρισης κειμένου χρησιμοποιεί μια σχεσιακή βάση δεδομένων για την αποθήκευση των χρηστών και των άρθρων. Για το σχεδιασμό της βάσης, κατασκευάστηκε ως πρώτο βήμα το μοντέλο οντοτήτων συσχετίσεων. Το συγκεκριμένο μοντέλο αποτελεί μια αφηρημένη αναπαράσταση των δεδομένων που πρόκειται να αποθηκευτούν στην βάση και χρησιμοποιείται κατά τη διάρκεια της ανάλυσης απαιτήσεων έτσι ώστε να εντοπιστούν και να καταγραφούν οι απαιτήσεις αναπαράστασης των δεδομένων που διαχειρίζεται η αναπτυσσόμενη εφαρμογή λογισμικού. Ακολουθεί η ανάλυση των οντοτήτων και των συσχετίσεων του συστήματος.

3.3.1 Οντότητες

3.3.1.1 Χρήστης

Η οντότητα χρήστης καθορίζει την πρόσβαση στις διάφορες λειτουργίες της εφαρμογής. Οι συνδεδεμένοι χρήστες - συντάκτες έχουν την δυνατότητα δημοσίευσης και επεξεργασίας άρθρων. Τα κυριότερα χαρακτηριστικά της οντότητας είναι το μοναδικό αναγνωριστικό, το όνομα χρήστη, ο κωδικός, το email, η ημερομηνία εγγραφής, η ημερομηνία τελευταίας σύνδεσης, η ημερομηνία αιτήματος ανάκτησης κωδικού.

3.3.1.2 Ομάδα

Η οντότητα ομάδα αναπαριστά τον λογικό διαχωρισμό των χρηστών σε ομάδες. Η κάθε ομάδα έχει ξεχωριστά προνόμια και πρόσβαση στις λειτουργίες του συστήματος. Παρατηρώντας το διάγραμμα χρήσης, προκύπτει εύκολα ο αριθμός των απαιτούμενων ομάδων της εφαρμογής.

3.3.1.3 Άρθρο

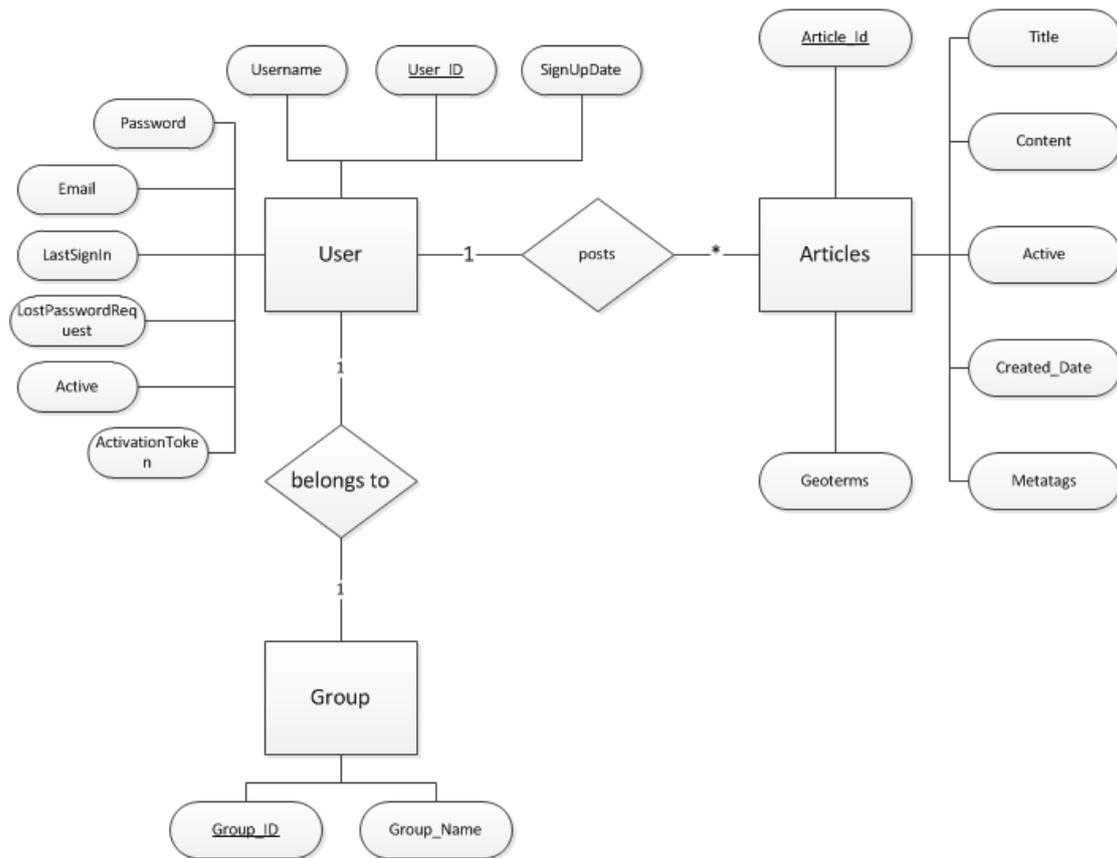
Η οντότητα άρθρο ενθυλακώνει όλα εκείνα τα στοιχεία που συνθέτουν μια δημοσίευση του συντάκτη. Περιλαμβάνει χαρακτηριστικά όπως είναι ο τίτλος, οι ετικέτες και το περιεχόμενο. Το πεδίο geoterms χρησιμοποιείται για την αποθήκευση της γεωγραφικής πληροφορίας που εντοπίζεται στο περιεχόμενο του άρθρου.

3.3.2 Συσχετίσεις

Στην εφαρμογή μας εμφανίζονται δύο συσχετίσεις ανάμεσα στις παραπάνω οντότητες. Η πρώτη σχέση είναι μεταξύ χρηστών και ομάδων. Κάθε χρήστης μπορεί να είναι

μέλος μιας ομάδας συνεπώς η σχέση είναι ένα προς ένα. Η δεύτερη σχέση είναι μεταξύ χρηστών και άρθρων. Κάθε χρήστης μπορεί να δημοσιεύσει πολλά άρθρα, επομένως η σχέση είναι ένα προς πολλά. Ακολουθεί το μοντέλο Οντοτήτων Συσχετίσεων που προκύπτει από την προηγούμενη ανάλυση.

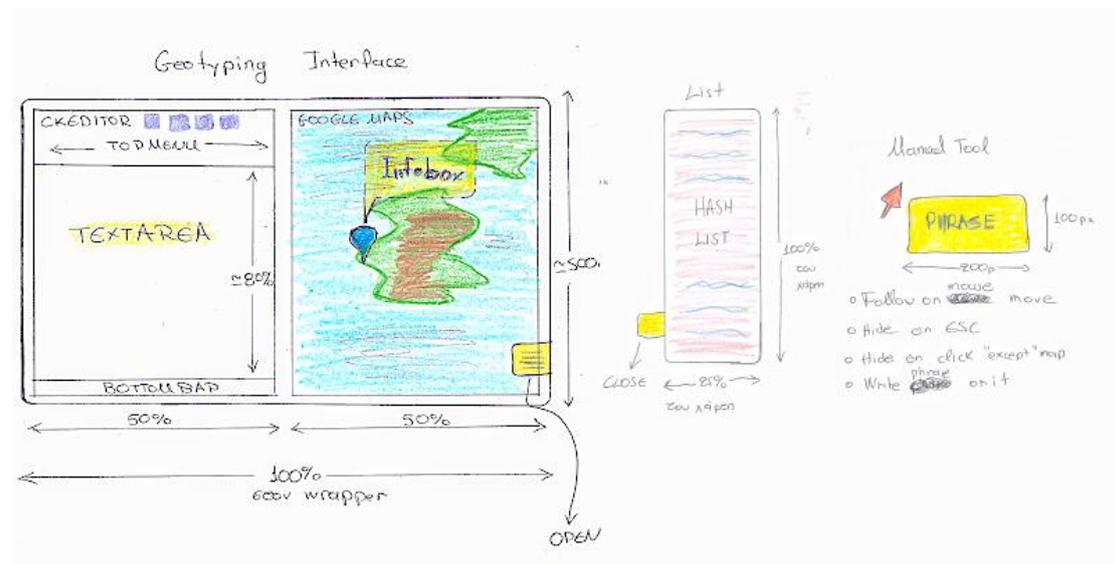
Διάγραμμα Οντοτήτων-Συσχετίσεων



3-4 Διάγραμμα Οντοτήτων-Συσχετίσεων

4. Σχεδίαση Συστήματος Geotyping

Στο παρόν κεφάλαιο αναλύεται η σχεδίαση της εφαρμογής Geotyping. Αρχικά περιγράφεται η αρχιτεκτονική του συστήματος ενώ στην συνέχεια παρουσιάζεται ο τρόπος λειτουργίας και η διασύνδεση των υποσυστημάτων.



4-1 Προσχέδιο γραφικού περιβάλλοντος

4.1 Αρχιτεκτονική

Το Geotyping είναι μια Ajax client-server εφαρμογή γραμμένη σε γλώσσα JavaScript. Για την υλοποίηση της εφαρμογής χρησιμοποιήθηκε η δημοφιλής βιβλιοθήκη jQuery, το CKEditor API, το Google Maps API και φυσικά επιστρατεύθηκαν οι τεχνολογίες HTML 5 και CSS 3. Τα υποσυστήματα δημιουργήθηκαν με κριτήριο τις διακριτές λειτουργικότητες της εφαρμογής. Ο proxy server που περιλαμβάνεται στην εφαρμογή γράφτηκε σε γλώσσα Php και σε Java μέσω της τεχνολογίας Java Server Pages (Jsp). Ο κώδικας οργανώνεται στους παρακάτω φακέλους και αρχεία:

Φάκελοι:

- scripts : Περιέχει όλα τα αρχεία JavaScript της εφαρμογής
 - ckeditor : Περιέχει τα αρχεία JavaScript του CKEditor 3.5.
 - geotyping : Περιλαμβάνει τα αρχεία JavaScript της εφαρμογής Geotyping.
- images : Φάκελος για τις εικόνες της εφαρμογής.
- styles : Φάκελος που περιλαμβάνει τα CSS (Cascading Styled Sheets) αρχεία που καθορίζουν το γραφικό περιβάλλον της εφαρμογής.

Αρχεία:

- index.html : Η βασική σελίδα της εφαρμογής.
- proxy_get.php : Proxy server σε Php μέσω GET αιτημάτων.
- proxy_post.php : Proxy server σε Php μέσω POST αιτημάτων.
- proxy_get.jsp : Proxy server σε Jsp μέσω GET αιτημάτων.

Στις ενότητες που ακολουθούν περιγράφουμε λεπτομερώς τα υποσυστήματα της εφαρμογής και τον τρόπο λειτουργίας τους.

4.2 Περιγραφή Κλάσεων

4.2.1 Geotyping.js

Η συγκεκριμένη κλάση-υποσύστημα αποτελεί τον βασικό κορμό της εφαρμογής. Αναλαμβάνει την αρχικοποίηση της εφαρμογής κατά την εκκίνηση, της διαχείρισης της ροής των δεδομένων και την εξαγωγή τους για μια ενδεχόμενη αποθήκευση. Οι μέθοδοι που περιλαμβάνει είναι οι εξής:

1. **Αρχικοποίηση ρυθμίσεων και εκκίνηση εφαρμογής.**
2. **Επεξεργασία της προηγούμενης παραγράφου του κείμενο που εισάγει ο χρήστης.**

Ο CKEditor είναι ένας HTML editor συνεπώς κάθε κείμενο που πληκτρολογείται ενθυλακώνεται στους κατάλληλους HTML κόμβους. Όλοι οι κόμβοι του κειμένου συνθέτουν ένα ιεραρχικό δέντρο ακολουθώντας του κανόνες που ορίζονται από την γλώσσα HTML.

Οι κόμβοι-στοιχεία στην HTML χωρίζονται σε δύο βασικές κατηγορίες: Τα **block elements** και τα **inline elements**. Τα block elements είναι ορθογώνια στοιχεία που δεν διαχωρίζονται σε γραμμές και μπορούν να έχουν ιδιότητες όπως είναι το πλάτος και το ύψος ανεξαρτήτων των γειτονικών στοιχείων. Στην κατηγορία ανήκουν για παράδειγμα τα στοιχεία <div>, <p>, <h1>, κτλ. Αντίθετα, τα inline elements ακολουθούν την ροή του κειμένου και μπορούν να διαχωριστούν σε γραμμές. Παραδείγματα inline elements αποτελούν τα στοιχεία <a>, , <script>, <code> και άλλα.

Λαμβάνοντας υπόψη τους παραπάνω ορισμούς θα αναλύσουμε τον τρόπο λειτουργίας του CKEditor. Το κείμενο που πληκτρολογεί ο χρήστης αποτελεί ουσιαστικά ένα inline element ενός block element, στην συγκεκριμένη περίπτωση ενός paragraph element.

Όπως αναφέρθηκε παραπάνω, το Geotyping λειτουργεί σε επίπεδο προηγούμενης παραγράφου. Με το όρο αυτό ομαδοποιούμε τα block elements <p>, <div>, <h1>, <h2>, <h3>, <h4>, <h5>, <h6>, <address> <pre> και εξάγουμε το κείμενο τους όταν αλλάζει το επιλεγμένο στοιχείο στον CKEditor.

Η μέθοδος επεξεργασία της προηγούμενης παραγράφου καλείται κάθε φορά που εγείρεται το γεγονός αλλαγής στοιχείου στον CKEditor είτε μέσω του πληκτρολογίου είτε μέσω του ποντικιού. Το υποσύστημα αναλαμβάνει να εξετάσει εάν το τρέχων και το προηγούμενο HTML στοιχείο ανήκει στην παραπάνω ομάδα

των block elements και εν συνεχεία εάν ανήκει, να ξεκινήσει την επεξεργασία του προηγούμενου αυτού στοιχείου για την εξαγωγή γεωγραφικής πληροφορίας.

Στο σημείο αυτό πρέπει να εξεταστούν δύο διαφορετικές περιπτώσεις εναλλαγής HTML στοιχείων στον CKEditor. Η πρώτη περίπτωση περιλαμβάνει την εισαγωγή μιας νέας παραγράφου. Ο χρήστης αρχικά πληκτρολογεί το κείμενο του και αλλάζει παράγραφο πατώντας το πλήκτρο Enter. Τότε ενεργοποιείται η μέθοδος και αναγνωρίζει ότι η προηγούμενη παράγραφος δεν είχε δεχθεί επεξεργασία στο παρελθόν ενώ η τρέχουσα παράγραφος είναι κενή. Σε αυτήν την περίπτωση το κείμενο της προηγούμενης παραγράφου προωθείται στον Geoparser.

Η δεύτερη περίπτωση περιλαμβάνει την διόρθωση μια παραγράφου που έχει δεχθεί επεξεργασία στο παρελθόν και περιέχει μαρκαρισμένες φράσεις. Ο χρήστης επιλέγει την συγκεκριμένη παράγραφο και την μεταβάλλει. Τότε το σύστημα αφαιρεί όλους τους μαρκαρισμένους όρους που περιέχει και τους αποθηκεύει σε μια προσωρινή δομή. Μετά την ολοκλήρωση της επεξεργασίας, ο χρήστης επιλέγει μια διαφορετική παράγραφο. Το σύστημα αναλαμβάνει να επαναφέρει όλους τους μαρκαρισμένους όρους που εξακολουθούν να υπάρχουν στο κείμενο μετά την επεξεργασία και στέλνει μόνο το τροποποιημένο κείμενο στον Geoparser. Οι όροι που δεν εντοπίζονται πλέον στην παράγραφο αφαιρούνται από τον χάρτη εάν εμφανίζονταν μόνο στην συγκεκριμένη παράγραφο. Η παραπάνω τεχνική προσωρινής αποθήκευσης στοχεύει στην μείωση του μεγέθους των αιτημάτων προς τον Geoparser και λειτουργεί μέσω ενός μηχανισμού μαρκαρίσματος που αναλύεται σε επόμενο κεφάλαιο.

3. Αποστολή αιτημάτων κειμένου στον proxy server του Geoparser.

Η μέθοδος εκτελεί ασύγχρονα Ajax αιτήματα και επεξεργάζεται τα αποτελέσματα που επιστρέφει ο Geoparser. Η απάντηση που λαμβάνει είναι σε XML μορφή οπότε η μέθοδος αναλαμβάνει να εξάγει τους πιθανούς γεωγραφικούς όρους από το XML και να τους προωθήσει στην μέθοδο επεξεργασίας γεωγραφικών όρων. Στην περίπτωση που χρησιμοποιείται ο εσωτερικός parser της εφαρμογής γίνεται απλά μια κλήση στην αντίστοιχη μέθοδο αυτού συντακτικού αναλυτή.

4. Επεξεργασία γεωγραφικών όρων που επιστρέφονται από τον Geoparser.

Κάθε πιθανός γεωγραφικός όρος, προωθείται στον Geocoder ώστε να βρεθούν οι συντεταγμένες του. Εάν ο όρος έχει ήδη εντοπιστεί σε προηγούμενο αίτημα, δηλαδή βρίσκεται στην δομή Geotyping Hash, τότε απλά αυξάνεται ο αριθμός εμφάνισης του στον χάρτη χωρίς περαιτέρω επεξεργασία. Σε κάθε όρο αντιστοιχεί ένας δείκτης στον χάρτη. Σε διαφορετική περίπτωση θα είχαμε

πολλαπλούς δείκτες στις ίδιες συντεταγμένες στον χάρτη, κάτι που δεν είναι επιθυμητό.

5. Έλεγχος για έγκυρη επιλογή κατά την χειροκίνητη τοποθέτηση φράσης στον χάρτη.

Η μέθοδος εξασφαλίζει ότι μόνο μη μαρκαρισμένο τμήμα του κείμενο που βρίσκεται σε μία γραμμή μπορεί να επιλεγεί από τον χρήστη για την χειροκίνητη αντιστοίχιση του σε συντεταγμένες στον χάρτη. Η επιλογή του χρήστη γίνεται με drag and drop μέσω του ποντικιού.

6. Εξαγωγή γεωγραφικών δεδομένων

Η μέθοδος αναλαμβάνει να εξάγει τους γεωγραφικούς όρους με τις συντεταγμένες τους από την δομή Geotyping Hash σε μορφή XML ή JSON Object. Το αποτέλεσμα της εξαγωγής μπορεί να αποθηκευτεί σε αρχείο στον δίσκο ή σε βάση δεδομένων.

4.2.2 Googlemaps.js

Η κλάση διαχειρίζεται όλο το κομμάτι που σχετίζει με του χάρτες και την γεωκωδικοποίηση. Οι μέθοδοι που περιλαμβάνει είναι οι ακόλουθες:

1. Εκκίνηση γραφικού περιβάλλοντος χάρτη και κουμπιών αλληλεπίδρασης

2. Γεωκωδικοποίηση, Geocoding

Η μέθοδος δέχεται ως είσοδο έναν γεωγραφικό όρο, επιστρέφει τις συντεταγμένες του και τοποθετεί έναν marker στον χάρτη στις συγκεκριμένες συντεταγμένες. Υποστηρίζεται ακόμη η μαζική γεωκωδικοποίηση όρων αποθηκευμένων σε ουρά προτεραιότητας. Η προσθήκη της συγκεκριμένης ουράς αναμονής αναλύεται σε επόμενο κεφάλαιο.

3. Αντίστροφη γεωκωδικοποίηση, Reverse Geocoding

Η μέθοδος δέχεται σαν είσοδο τις συντεταγμένες ενός σημείου στον χάρτη και επιστρέφει την διεύθυνση του αντίστοιχου σημείου. Η λειτουργία χρησιμοποιείται για την χειροκίνητη διόρθωση των δεικτών στον χάρτη.

4. Δημιουργία και απεικόνιση δείκτη στον χάρτη

Η μέθοδος δημιουργεί ένα δείκτη (marker) και ένα παράθυρο πληροφοριών (info window) που συνδέεται με αυτό. Ο δείκτης αποθηκεύεται στην δομή Geotyping Hash μαζί με τους αντίστοιχους γεωγραφικούς όρους του κειμένου.

5. Διαγραφή marker

Ο χρήστης έχει την δυνατότητα να διαγράψει λανθασμένους δείκτες που τοποθετούνται στον χάρτη από τον Geocoder.

4.2.3 *Gtm.js*

Η κλάση περιλαμβάνει τον κώδικα που απαιτείται για την χειροκίνητη τοποθέτηση ενός όρου στον χάρτη. Περιλαμβάνει τις παρακάτω μεθόδους:

1. **Listeners για τα γεγονότα του ποντικιού**

Η κλάση διαχειρίζεται γεγονότα κίνησης και κλικ του ποντικιού σε περιοχές εντός ή εκτός του χάρτη τα οποία είναι απαιτούνται για την εκτέλεση της λειτουργίας.

2. **Δημιουργία και εμφάνιση του πλαισίου χειροκίνητης αντιστοίχισης**

Το πλαίσιο χειροκίνητης απεικόνισης σηματοδοτεί την έναρξη της λειτουργίας χειροκίνητης αντιστοίχισης. Το πλαίσιο εμφανίζεται μετά την έγκυρη επιλογή μιας φράσης στο κείμενο. Στόχος του χρήστη είναι η τοποθέτηση του πλαισίου στο επιθυμητό σημείο στον χάρτη.

4.2.4 *Hash.js*

Η κλάση περιέχει την δομή δεδομένων (Geotyping Hash) στην οποία αποθηκεύονται οι γεωγραφικοί όροι του κειμένου με τους αντίστοιχους δείκτες. Πρόκειται για ένα hash table που έχει δυνατότητες ταυτόχρονης γραφικής απεικόνισης των περιεχομένων του με την μορφή HTML λίστας. Οι μέθοδοι που περιλαμβάνονται στην κλάση είναι οι εξής:

1. **Εισαγωγή, Επεξεργασία και διαγραφή στοιχείου από τη δομή**

Οι μέθοδοι αποτελούν την διεπαφή της δομής με την εφαρμογή.

2. **Εκτύπωση όλων των περιεχομένων της δομής**

Εξαγωγή όλων των περιεχομένων της δομής σε μορφή κειμένου.

3. **Εμφάνιση της HTML λίστας της δομής**

Γραφική παρουσίαση των περιεχομένων της δομής για καλύτερη εποπτεία.

4.2.5 *Highlight.js*

Πρόκειται για την επέκταση του jQuery Highlight Plug-in του Johann Burkard. Αναλαμβάνει να επισημάνει στο κείμενο τους γεωγραφικούς όρους που τοποθετήθηκαν στο χάρτη. Η επισήμανση γίνεται με κίτρινο φόντο σε κάθε όρο. Πριν την μεταβολή του κειμένου από τον χρήστη το υποσύστημα αναλαμβάνει να αφαιρέσει την επισήμανση των όρων της παραγράφου.

4.2.6 *Parser.js*

Το υποσύστημα είναι ένας απλοϊκός εσωτερικός συντακτικός αναλυτικής γραμμένος σε JavaScript που δέχεται ως είσοδο κείμενο, το διαχωρίζει σε προτάσεις, λέξεις και επιστρέφει λέξεις οι ακολουθίες λέξεων που ξεκινούν με κεφαλαίο γράμμα. Χρησιμοποιήθηκε αρκετά κατά την διάρκεια της ανάπτυξης της εφαρμογής καθώς πλεονεκτεί σε ταχύτητα σε σχέση με τον εξωτερικό parser.

4.2.7 *XmlWrite.js*

Πρόκειται για μια JavaScript βιβλιοθήκη η οποία αναλαμβάνει να παράγει XML αρχεία. Χρησιμοποιείται για την εξαγωγή των γεωγραφικών δεδομένων από την δομή Geotyping Hash σε XML. Το εικονικό αρχείο που παράγεται μπορεί να αποθηκευτεί στον δίσκο ή σε βάση δεδομένων.

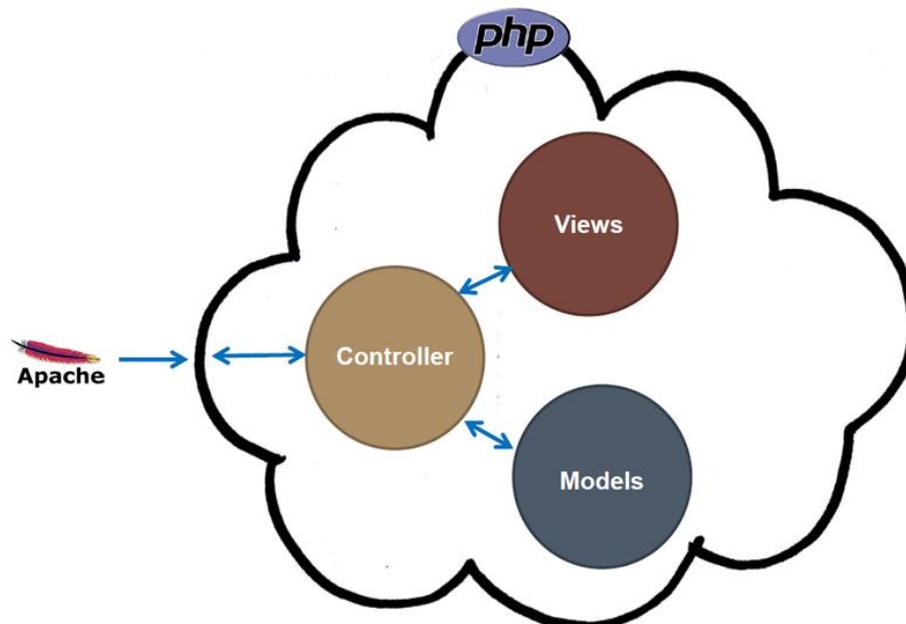
5. Σχεδίαση συστήματος *Geotyping Text*

Management System

Στο παρόν κεφάλαιο αναλύεται η υλοποίηση του *Geotyping Text Management System*. Αρχικά περιγράφεται η αρχιτεκτονική του συστήματος και στην συνέχεια δίνονται οι κλάσεις της εφαρμογής με τις λειτουργίες που εκτελεί η κάθε μια.

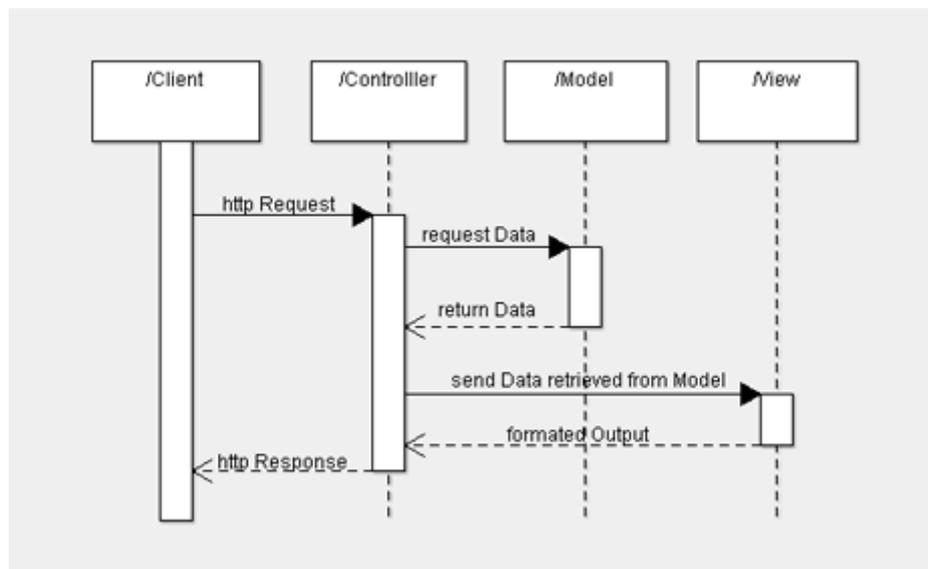
5.1 Αρχιτεκτονική

Για την υλοποίηση της εφαρμογής επιλέχθηκε το αρχιτεκτονικό πρότυπο Model View Controller (MVC). Μία εφαρμογή που αναπτύσσεται ακολουθώντας τη συγκεκριμένη τεχνοτροπία απαρτίζεται από ξεχωριστές δομικές μονάδες που υλοποιούν τους τρεις διακριτούς ρόλους που ορίζονται από τη συγκεκριμένη αρχιτεκτονική.



5-1 Μοντέλο MVC με PHP και Apache Server

Η τεχνοτροπία αυτή βασίζεται στη διάσπαση ενός συστήματος σε τρεις μονάδες, τη μονάδα *Μοντέλο (Model)* που περιλαμβάνει και μοντελοποιεί τις βασικές λειτουργίες και τα δεδομένα του συστήματος, μια ή περισσότερες μονάδες *Προβολής (View)* που ορίζουν διαφορετικούς τρόπους παρουσίασης δεδομένων και αποτελεσμάτων στο χρήστη και τέλος μια ή περισσότερες μονάδες *Ελέγχου (Controller)* που διαχειρίζονται με τον κατάλληλο τρόπο την εισαγωγή πληροφοριών και αιτημάτων από τον χρήστη. Στην συνέχεια εμβαθύνουμε στις τρεις βασικές μονάδες του μοντέλου MVC. Στο ακολουθιακό διάγραμμα του σχήματος 8 φαίνεται ο τρόπος που αλληλεπιδρούν οι μονάδες μεταξύ τους. Για την υλοποίηση του συστήματος επιλέχθηκε η γλώσσα PHP.



5-2 Ακολουθιακό διάγραμμα MVC προτύπου

5.1.1 Model

Αποτελεί την αναπαράσταση των δεδομένων που διαχειρίζεται η εφαρμογή καθώς και τις βασικές λειτουργίες που επιδρούν στα δεδομένα αυτά. Τα μοντέλα αποτελούν μια αφαιρετική βαθμίδα μεταξύ της εφαρμογής και του μόνιμου μηχανισμού αποθήκευσης. Συχνά ενθυλακώνουν τους λειτουργικούς κανόνες που θέτει η κάθε εφαρμογή (Business Logic). Σε κάθε μεταβολή του μοντέλου, ενημερώνονται οι αντίστοιχες προβολές ώστε να η αλλαγή αυτή να γίνει ορατή στον τελικό χρήστη.

5.1.2 View

Η προβολή είναι αποκλειστικά υπεύθυνη για την παραγωγή της διεπαφής της εφαρμογής. Προβάλλει ένα μοντέλο σε μορφή κατανοητή από το χρήστη και προσαρμοσμένη στο μέσο απεικόνισης. Συνήθως, παρέχονται διαφορετικά views για κάθε μοντέλο τα οποία χρησιμοποιούνται ανάλογα την εκάστοτε ανάγκη παρουσίασης της πληροφορίας του

μοντέλου. Συνηθισμένο παράδειγμα είναι η ύπαρξη μιας προβολής με συνοπτικές πληροφορίες για το μοντέλο και μιας δεύτερης προβολής, που απεικονίζει αναλυτικά όλα τα στοιχεία του μοντέλου (master detail view).

Εκτός από τη διάκριση στον τρόπο με τον οποίο απεικονίζεται το μοντέλο, τα views διακρίνονται και με βάση το μέσο στο οποίο είναι σχεδιασμένα να προβάλλονται. Πολλές φορές ένα μεγάλο σύστημα παρέχει στους χρήστες του πολλαπλές δυνατότητες πρόσβασης και απεικόνισης της πληροφορίας, π.χ. μέσω του web browser, μέσω μίας ανεξάρτητης γραφικής εφαρμογής, μέσω μίας ανεξάρτητης εφαρμογής κειμένου ή μέσω εκτυπωμένης πληροφορίας. Σε κάθε διαφορετικό μέσο απεικόνισης αντιστοιχεί και ένα διαφορετικό view, το οποίο λαμβάνει υπόψη του τις ιδιαιτερότητες του μέσου ώστε να εμφανίσει με το βέλτιστο δυνατό τρόπο τα στοιχεία του μοντέλου, π.χ. η προβολή ενός γραφήματος στην κονσόλα του συστήματος ή η προβολή μίας έγχρωμης εικόνας σε μία ασπρόμαυρη εκτύπωση είναι περιπτώσεις που απαιτούν ιδιαίτερο χειρισμό από το αντίστοιχο view.

Το σύστημα που μελετάμε είναι μια δικτυακή εφαρμογή. Συνεπώς οι προβολές περιλαμβάνουν HTML κώδικα και τον απαραίτητο κώδικα PHP. Οι προβολές χρησιμοποιούν μηχανισμούς προτύπων (templates) που τους επιτρέπουν πρόσβαση μόνο σε συγκεκριμένες ονομασμένες μεταβλητές πετυχαίνοντας τον βέλτιστο διαχωρισμό διεπαφής χρήστη και δεδομένων.

5.1.3 Controller

Λαμβάνει τα μηνύματα και τις εντολές εισόδου από το χρήστη και αποκρίνεται ενεργώντας πάνω στα μοντέλα ή/και εμφανίζοντας κάποια συγκεκριμένη προβολή. Ο ελεγκτής αποτελεί το συνδετικό κρίκο του συστήματος και είναι υπεύθυνος για τη διεκπεραίωση των εντολών του χρήστη της εφαρμογής. Η συνήθης λειτουργία του είναι η αναγνώριση των εντολών του χρήστη, η άντληση και επεξεργασία της πληροφορίας από τα μοντέλα βάσει των εντολών του χρήστη και η παρουσίαση της ζητούμενης πληροφορίας μέσω του κατάλληλου view.

5.2 Οργάνωση φακέλων και αρχείων

Για την υλοποίηση του συστήματος διαχείρισης κειμένων χρησιμοποιήθηκε στην πλευρά του εξυπηρετητή η γλώσσα PHP 5 ενώ στην πλευρά του πελάτη οι τεχνολογίες HTML 5, CSS 3 και η γλώσσα JavaScript μέσω της βιβλιοθήκης jQuery. Ο κώδικας οργανώνεται στους παρακάτω φακέλους και αρχεία:

Φάκελοι:

- application: Περιέχει τις γενικές κλάσεις της εφαρμογής.
- controllers: Περιλαμβάνονται οι controllers που επεξεργάζονται τα αιτήματα των χρηστών.
- css: Περιλαμβάνει τα Cascading Styled Sheets αρχεία που καθορίζουν την διεπαφή του συστήματος
- email_templates: Φάκελος που περιέχει τα πρότυπα emails προς αποστολή.
- images: Περιέχει τις απαιτούμενες εικόνες της εφαρμογής.
- includes: Περιέχει βιβλιοθήκες χρήσιμων εργαλείων σε PHP.
- install: Ο φάκελος πρώτης εγκατάστασης της εφαρμογής. Μετά την επιτυχή εγκατάσταση ο φάκελος πρέπει να διαγραφεί.
- lang: Φάκελος που χρησιμοποιείται για το localization της εφαρμογής. Περιέχει τα αρχεία που ορίζουν το περιεχόμενο μηνυμάτων και ετικετών σε κάθε υποστηριζόμενη γλώσσα.
- lib: Περιέχει βιβλιοθήκες συναρτήσεων σε PHP.
- models: Περιλαμβάνονται οι κλάσεις των μοντέλων του προτύπου MVC.
- modules: Ο φάκελος περιλαμβάνει επαναχρησιμοποιούμενες ενότητες των προβολών
- scripts: Περιέχει τα JavaScript αρχεία συμπεριλαμβανομένης της εφαρμογής Geotyping.
- views: Περιλαμβάνονται οι προβολές του μοντέλου MVC.

Αρχεία:

- index.php: Αποτελεί το μοναδικό σημείο πρόσβασης στο σύστημα όπως αναλύεται παρακάτω.

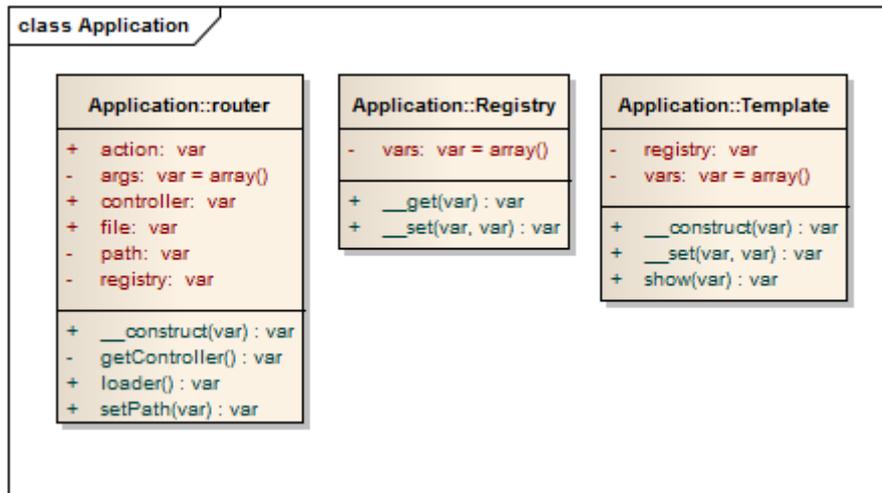
5.3 Βασικά χαρακτηριστικά

Τα βασικά χαρακτηριστικά της εφαρμογής συνοψίζονται στα παρακάτω σημεία:

- Μοναδικό σημείο εισόδου στο σύστημα είναι το αρχείο index.php από το οποίο ελέγχεται η πρόσβαση στο σύστημα. Ανάλογα με την διεύθυνση (URL) που ζητάει ο χρήστης το αρχείο index επιλέγει τον κατάλληλο controller για να επεξεργαστεί το αίτημα μέσω ενός διαβιβαστή (router). Η διαδικασία αναλύεται σε επόμενο κεφάλαιο.
- Δυναμική φόρτωση των κλάσεων που απαιτούνται σε κάθε αίτημα μέσω της συνάρτησης autoload που προσφέρει η PHP.
- Υποστήριξη προτύπων (templating) στις προβολές για αποδοτικότερο διαχωρισμό HTML και PHP κώδικα.
- Δυνατότητα localization για προσθήκη πολλαπλών γλωσσών.
- Υποστήριξη φιλικών διευθύνσεων προς τον χρήστη και τις μηχανές αναζήτησης (SEF Search engines friendly URLs).
- Ενσωμάτωση ετικετών στα άρθρα για εύκολη αναζήτηση και κατηγοριοποίηση του περιεχομένου.

5.4 Περιγραφή κλάσεων

5.4.1 Γενικές κλάσεις εφαρμογής



5-3 Διάγραμμα γενικών κλάσεων εφαρμογής

5.4.1.1 registry.class.php

Είναι το αντικείμενο στο οποίο αποθηκεύονται οι μεταβλητές που απαιτούνται σε κάθε αίτημα του χρήστη και χρησιμοποιείται από τον controller. Προστατεύει το σύστημα από την υπερβολική χρήση καθολικών μεταβλητών καθώς ο κάθε ελεγκτής διαθέτει ένα registry αντικείμενο με ακριβώς όσα δεδομένα χρειάζεται.

Ιδιότητες:

- vars: Σχεσιακός πίνακας στον οποίο αποθηκεύονται οι πληροφορίες που δίνονται μέσω των get & set.

Μέθοδοι:

- `__get()`: Ανάκτηση πληροφορίας από το αντικείμενο vars.
- `__set()`: Αποθήκευση πληροφορίας στο αντικείμενο vars.

5.4.1.2 router.class.php

Αποτελεί τον διαβιβαστή (*dispatcher*) του συστήματος ο οποίος ελέγχει το URL και αποστέλλει την αίτηση στον κατάλληλο ελεγκτή.

Ιδιότητες:

- file: Αρχείο ελεγκτή που πρέπει να φορτωθεί.

- `controller`: Όνομα κλάσης ελεγκτή.
- `action`: Το όνομα της μέθοδος του επιλεγμένου ελεγκτή που θα αναλάβει την επεξεργασία του αιτήματος.

Μέθοδοι:

- `loader()`: Εντοπίζει και φορτώνει τον αντίστοιχο ελεγκτή. Σε περίπτωση λάθους επιστρέφει 404 σφάλμα.
- `getController()`: Μέθοδος η οποία επεξεργάζεται το URL και επιστρέφει το όνομα του ελεγκτή που θα χρησιμοποιηθεί και την ενέργεια που αυτός θα εκτελέσει.

5.4.1.3 *template.class.php*

Αποτελεί χαρακτηριστικό της αρχείων προβολής δίνοντας στις προβολές πρόσβαση μόνο σε συγκεκριμένες μεταβλητές.

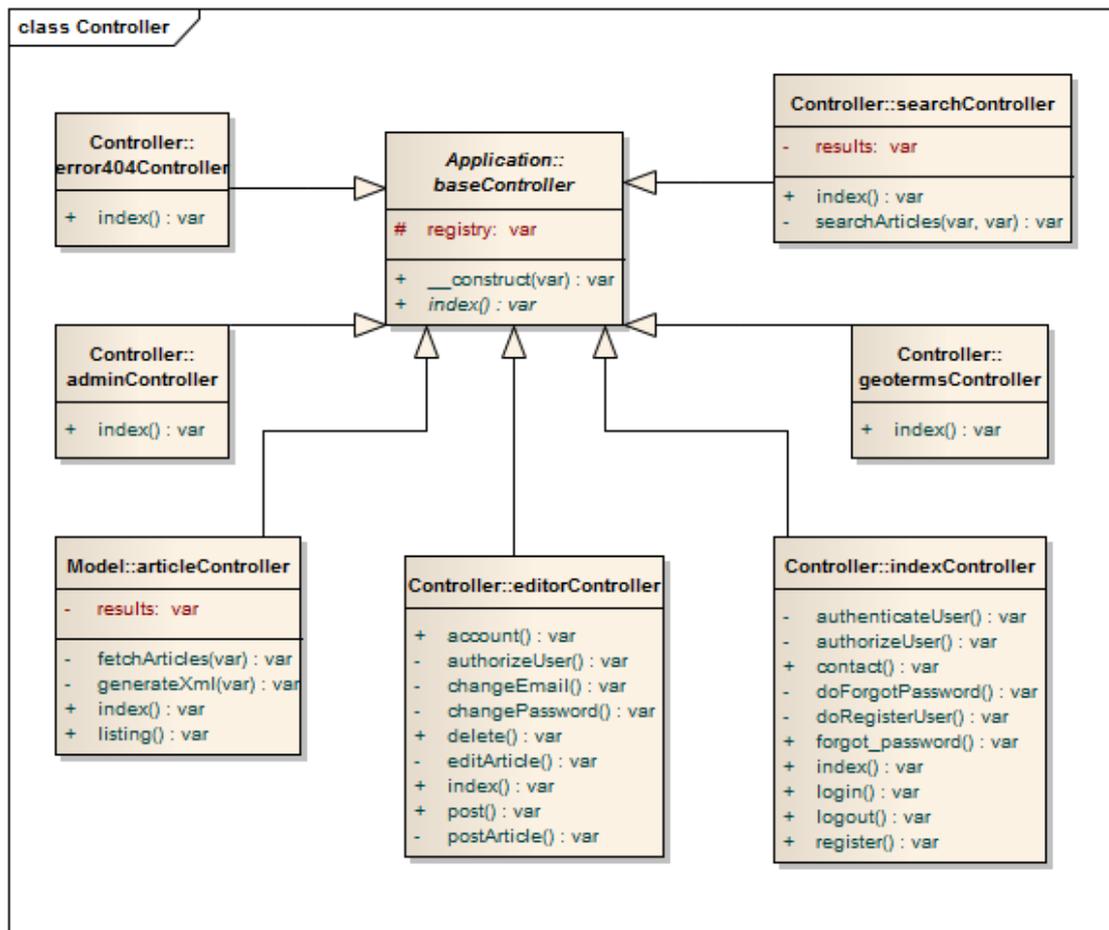
Ιδιότητες:

- `registry`: Αντικείμενο registry που αναλύθηκε παραπάνω.
- `vars`: Δομή αποθήκευσης μεταβλητών που θα προσπελαστούν από την εκάστοτε προβολή.

Μέθοδοι:

- `__get()`: Ανάκτηση πληροφορίες από το αντικείμενο vars.
- `__set()`: Αποθήκευση πληροφορίας στο αντικείμενο vars.
- `show()`: Αρχικοποιεί και φορτώνει το template κάνοντας διαθέσιμες στην αντίστοιχη προβολή, τις μεταβλητές που περιέχονται στην δομή vars.

5.4.2 Controllers



5-4 Διάγραμμα κλάσεων ελεγκτών

5.4.2.1 controller_base.class.php

Αφηρημένη υλοποίηση του αντικειμένου controller.

Ιδιότητες:

- registry: Αντικείμενο registry που ενθυλακώνεται σε κάθε ελεγκτή.

Μέθοδοι:

- index(): Κάθε controller υλοποιεί την μέθοδο index ως την προκαθορισμένη λειτουργία που πρέπει να εκτελεί.

5.4.2.2 adminController.php

Ελεγκτής που διαχειρίζεται τα αιτήματα του διαχειριστή του συστήματος.

5.4.2.3 *articleController.php*

Ελεγκτής που διαχειρίζεται τα αιτήματα προβολής ενός ή όλων των δημοσιευμένων άρθρων.

Μέθοδοι:

- `index()`: Ανάκτηση και προβολή ενός συγκεκριμένου άρθρου.
- `listing()`: Ανάκτηση και προβολή όλων των δημοσιευμένων άρθρων.

5.4.2.4 *editorController.php*

Διαχειρίζεται όλα τα αιτήματα του συνδεδεμένου χρήστη - συντάκτη.

Μέθοδοι:

- `index()`: Προβολή της αρχικής σελίδας του συντάκτη.
- `post()`: Δημοσίευση ή επεξεργασία ενός νέου άρθρου.
- `delete()`: Διαγραφή άρθρου.
- `account()`: Διαχείριση στοιχείων λογαριασμού χρήστη.

5.4.2.5 *error404.php*

Ελεγκτής που χρησιμοποιείται στην περίπτωση σφάλματος κατά την επεξεργασία του αιτήματος.

5.4.2.6 *geotermsController.php*

Αναλαμβάνει να επιστρέψει το XML με τους γεωγραφικούς όρους του κάθε άρθρου. Χρησιμοποιείται από τον χάρτη καθώς αρχικοποιείται η προβολή του άρθρου.

5.4.2.7 *indexController.php*

Ο αρχικός ελεγκτής που διαχειρίζεται τα αιτήματα του ανώνυμου χρήστη του συστήματος. Οι αιτήσεις για εγγραφή, σύνδεση και ανάκτηση κωδικού χρήστη εξυπηρετούνται από αυτόν τον ελεγκτή.

Μέθοδοι:

- `index()`: Επιστροφή αρχικής σελίδας εφαρμογής.
- `register()`: Εγγραφή χρήστη στο σύστημα.
- `login()`: Σύνδεση χρήστη.

- `logout()`: Αποσύνδεση χρήστη.
- `forgot_password()`: Διαχείριση αιτήματος ανάκτησης κωδικού
- `contact()`: Επεξεργασία και αποστολή φόρμας επικοινωνίας

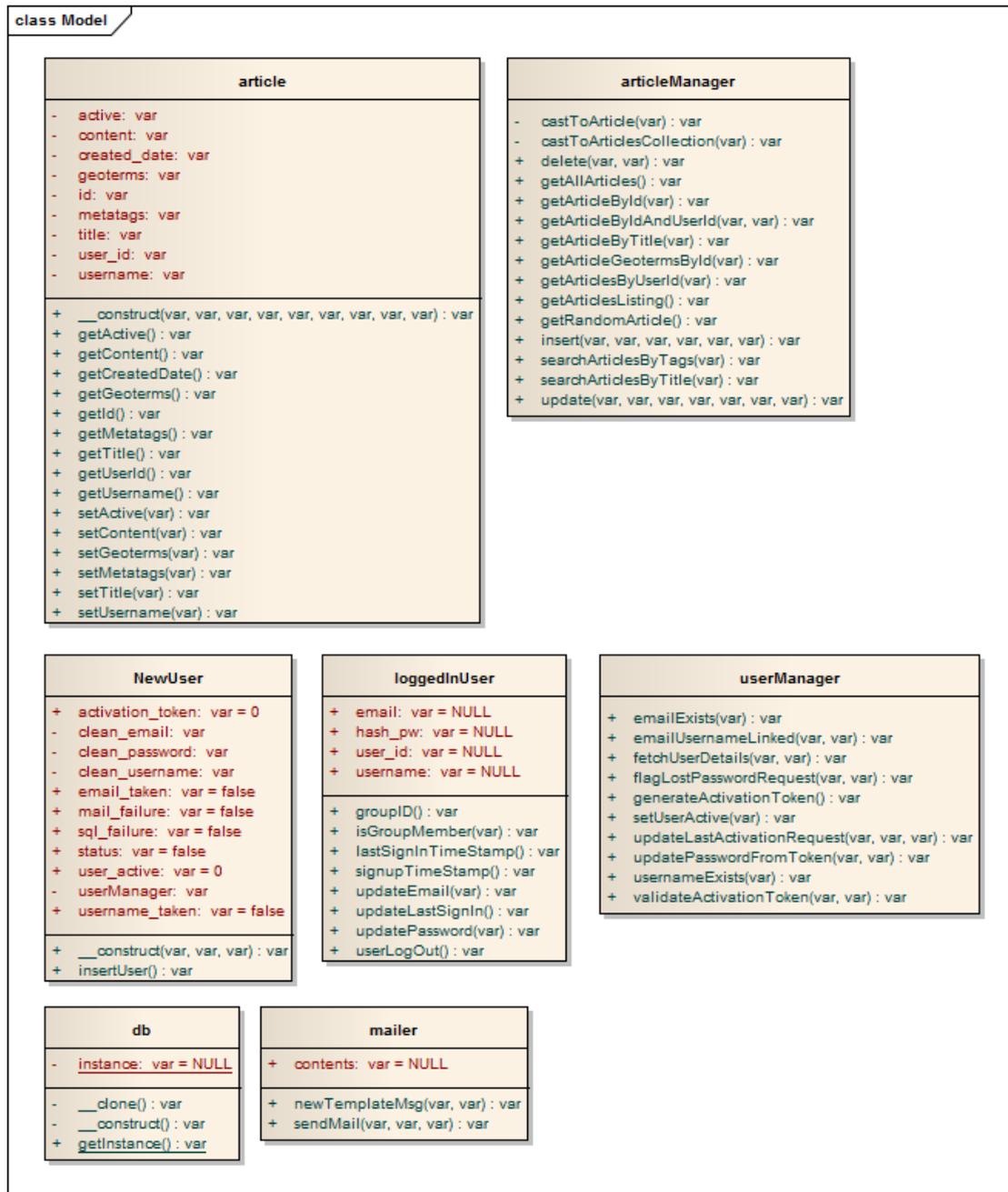
5.4.2.8 *searchController.php*

Διαχειρίζεται τα αιτήματα αναζήτησης άρθρων. Η αναζήτηση γίνεται είτε με βάση τις ετικέτες είτε με βάση τον τίτλο του άρθρου. Ο ελεγκτής, έχοντας ως παράμετρο την λέξη κλειδί που αναζητάει ο χρήστης, εκτελεί το κατάλληλο ερώτημα στην βάση δεδομένων μέσω του μοντέλου της εφαρμογής και επιστρέφει τα αποτελέσματα στην διεπαφή.

Μέθοδοι:

- `index()`: Αναζήτηση άρθρου

5.4.3 Models



5-5 Διάγραμμα κλάσεων μοντέλου

5.4.3.1 article.class.php

Αντικείμενο που αναπαριστά την οντότητα του άρθρου.

5.4.3.2 *articleManager.class.php*

Αντικείμενο που διαχειρίζεται άρθρα. Αναλαμβάνει την ανάκτηση, την αποθήκευση, την επεξεργασία και την διαγραφή άρθρων από την βάση. Στις κλάσεις αυτές περιλαμβάνονται τα SQL ερωτήματα που εκτελούνται στην βάση.

5.4.3.3 *db.class.php*

Πρόκειται για το αντικείμενο της σύνδεσης της εφαρμογής με την βάση δεδομένων. Το αντικείμενο εξασφαλίζει ότι ο κάθε χρήστης έχει στην κατοχή του μόνο μια σύνδεση προς την βάση την οποία και επαναχρησιμοποιεί. Συνεπώς βελτιστοποιείται η χρήση των πόρων του εξυπηρετητή της βάσης δεδομένων. Το αντικείμενο χρησιμοποιεί την βιβλιοθήκη MySQL 4 Database Abstraction Layer της PHP που αναλαμβάνει την επικοινωνία της εφαρμογής με την MySQL βάση, την διαχείριση σφαλμάτων που πιθανόν να προκύψουν και την προστασία από επιθέσεις τύπου SQL Injection.

5.4.3.4 *loggedInUser.class.php*

Κλάση που αναπαριστά τον συνδεδεμένο χρήστη στο σύστημα. Περιλαμβάνει όλες τις πληροφορίες που απαιτούνται από τις λειτουργίες του συντάκτη καθώς επίσης και μεθόδους για την αποσύνδεση και την επεξεργασία του προφίλ του χρήστη.

5.4.3.5 *mailer.class.php*

Αποτελεί το υποσύστημα δημιουργίας και αποστολής email.

5.4.3.6 *newUser.class.php*

Αντικείμενο που μοντελοποιεί την οντότητα του χρήστη.

5.4.3.7 *userManager.class.php*

Αντικείμενο που διαχειρίζεται τους χρήστες και αναλαμβάνει την καταχώρηση νέων χρηστών στο σύστημα κατά την διαδικασία εγγραφής τους.

5.4.4 **Views**

Οι προβολές περιλαμβάνουν HTML κώδικα και τον απαραίτητο PHP κώδικα για την παραγωγή της γραφικής διεπαφής χρήστη. Οι προβολές που υποστηρίζονται από το σύστημα είναι οι ακόλουθες:

5.4.4.1 *editor/account.php*

Προβολή της σελίδας διαχείρισης λογαριασμού του συντάκτη.

5.4.4.2 *editor/index.php*

Προβολή της αρχικής σελίδας του συντάκτη.

5.4.4.3 *editor/post.php*

Προβολή της σελίδας δημοσίευσης και επεξεργασίας άρθρων. Στην προβολή αυτή περιλαμβάνεται η εφαρμογή Geotyping.

5.4.4.4 *contact.php*

Προβολή φόρμας επικοινωνίας.

5.4.4.5 *error404.php*

Προβολή σε περίπτωση σφάλματος.

5.4.4.6 *forgot_password.php*

Προβολή υπενθύμισης κωδικού πρόσβασης.

5.4.4.7 *index.php*

Προβολή αρχικής σελίδας εφαρμογής

5.4.4.8 *listing.php*

Προβολή λίστας όλων των δημοσιευμένων άρθρων.

5.4.4.9 *login.php*

Προβολή σελίδας σύνδεσης χρήστη.

5.4.4.10 *register.php*

Προβολή σελίδας εγγραφής χρήστη.

5.4.5 *lib*

5.4.5.1 *funcs.general.php*

Βιβλιοθήκη με συναρτήσεις γενικού σκοπού.

5.4.5.2 *funcs.user.php*

Βιβλιοθήκη με συναρτήσεις για την διαχείριση και την σύνδεση χρηστών στο σύστημα.

6. Υλοποίηση

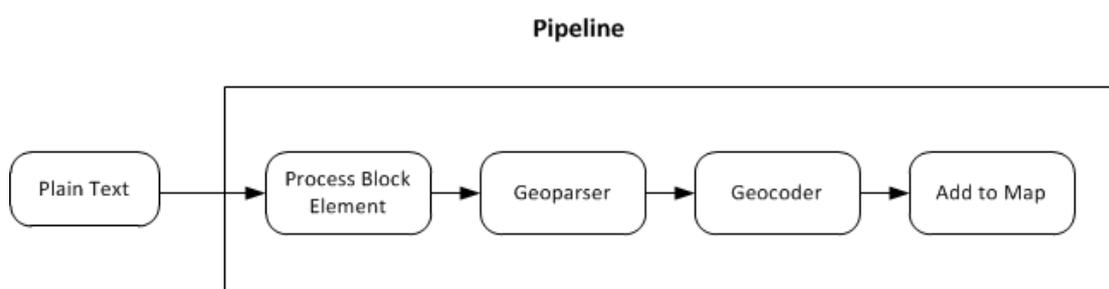
Στον παρόν κεφάλαιο περιγράφονται τεχνικές λεπτομέρειες της υλοποίησης των δύο συστημάτων που παρουσιάζουν ιδιαίτερο ενδιαφέρον. Αρχικά εστιάζουμε στα τεχνικά χαρακτηριστικά της εφαρμογής Geotyping. Έπειτα αναλύουμε μερικά από τα χαρακτηριστικά της αρχιτεκτονικής του συστήματος διαχείρισης κειμένου Geotyping Text Management System.

6.1 Λεπτομέρειες υλοποίησης της εφαρμογής Geotyping

6.1.1 Επισήμανση στοιχείων και Pipeline επεξεργασίας.

Στην παράγραφο που ακολουθεί περιγράφεται ο αλγόριθμος επεξεργασίας των παραγράφων του κειμένου. Για να απλοποιήσουμε την περιγραφή του αλγορίθμου, θεωρούμε ότι το κείμενο μας αποτελείται αποκλειστικά από παραγράφους.

Με βάση αυτή την παραδοχή, ο κύκλος επεξεργασίας της προηγούμενης παραγράφου του κειμένου ενεργοποιείται με την επιλογή μιας διαφορετικής παραγράφου. Ένα παράδειγμα αποτελεί το πάτημα του πλήκτρου enter που δημιουργεί και επιλέγει μια επόμενη, νέα παράγραφο. Η προηγούμενη παράγραφος δέχεται επεξεργασία από ένα pipeline αποτελούμενο από τις λειτουργίες του CKEditor, του Geoparser και του Geocoder.



6-1 Geotyping Pipeline

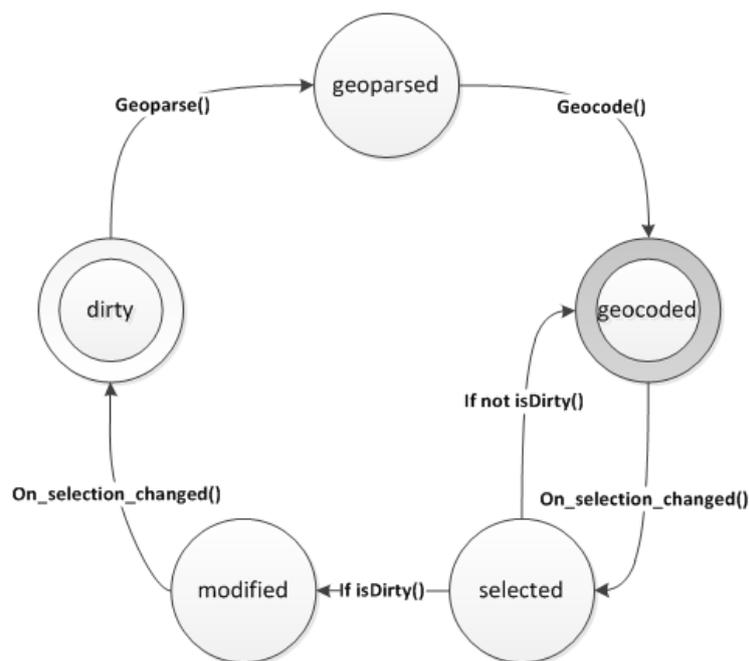
Η ορθή επεξεργασία του κειμένου από τις παραπάνω διαδικασίες οδήγησε στην ανάγκη ενός μηχανισμού επισήμανσης των παραγράφων ώστε να μπορεί το σύστημα να αναγνωρίζει την κατάσταση και την θέση της κάθε παραγράφου στο pipeline.

Ο μηχανισμός μαρκαρίσματος αποτελείται από πέντε καταστάσεις. Ως αρχική κατάσταση ορίζουμε την κατάσταση "dirty" η οποία σηματοδοτεί την εξαγωγή νέου κειμένου από τον CKEditor. Αφού το κείμενο της παραγράφου σταλεί στον Geoparser η παράγραφος μεταβαίνει στην κατάσταση "geoparsed". Στην συνέχεια, μετά την ολοκλήρωση της γεωκωδικοποίησης των όρων του κειμένου, η παράγραφος μεταβαίνει στην κατάσταση "geocoded". Η κατάσταση "geocoded" ορίζεται ως τελική κατάσταση.

Οι υπόλοιπες δύο καταστάσεις απαιτούνται για την διόρθωση μιας προηγούμενης παραγράφου. Μια παράγραφος που έχει δεχθεί επεξεργασία στο παρελθόν, βρίσκεται στην κατάσταση "geocoded". Εάν ο συντάκτης επιλέξει να την διορθώσει, η παράγραφος μεταβαίνει στην κατάσταση "selected" που σηματοδοτεί την πιθανή τροποποίηση της συγκεκριμένης παραγράφου. Εάν η παράγραφος μεταβληθεί τότε μεταβαίνει στην κατάσταση modified και ο κύκλος επεξεργασίας ξεκινά από την αρχή. Σε διαφορετική περίπτωση, η παράγραφος επανέρχεται στην τελική κατάσταση "geocoded" που ήταν προηγουμένως.

Ολοκληρώνοντας την περιγραφή του μηχανισμού αξίζει να αναφέρουμε το γεγονός ότι η κατάσταση modified επιτρέπει την επαναφορά των γεωγραφικών όρων που υπήρχαν στον κείμενο και εξακολουθούν να περιέχονται σε αυτό, ελαχιστοποιώντας το δεδομένα που πρέπει να ακολουθήσουν τον νέο κύκλο επεξεργασίας.

Διάγραμμα καταστάσεων μηχανισμού μαρκαρίσματος



6-2 Διάγραμμα καταστάσεων μηχανισμού μαρκαρίσματος

6.1.2 Proxy Server και Cross-Domain Ajax Request

Οι δημοφιλείς περιηγητές (browsers) δεν επιτρέπουν αιτήσεις Ajax μεταξύ διαφορετικών ιστοτόπων μέσω του same origin policy. Πρόκειται για μια σημαντική έννοια ασφάλειας για τις γλώσσες προγραμματισμού στην πλευρά του πελάτη, όπως είναι η JavaScript. Η πολιτική επιτρέπει σε τμήματα του κώδικα που τρέχουν σε σελίδες του ίδιου ιστότοπου, να αλληλεπιδρούν αλλά απαγορεύει την πρόσβαση σε μεθόδους και ιδιότητες σελίδων διαφορετικών ιστοτόπων. Οι αιτήσεις μεταξύ ιστοτόπων θεωρούνται απόπειρα επίθεσης cross site scripting (XSS).

Ο παραπάνω μηχανισμός έχει ιδιαίτερη σημασία για τις σύγχρονες διαδικτυακές εφαρμογές που σε μεγάλο βαθμό εξαρτώνται από τα HTTP cookies. Τα μπισκότα (cookies) συχνά χρησιμοποιούνται για τον έλεγχο της ταυτότητας του χρήστη μέσω του μηχανισμού των sessions. Οι διακομιστές από την πλευρά τους, ενεργούν με βάση τις πληροφορίες των μπισκότων για να αποκαλύψουν ευαίσθητες πληροφορίες ή να εκτελέσουν συγκεκριμένες λειτουργίες. Ο αυστηρός διαχωρισμός μεταξύ του περιεχομένου που παρέχεται από ξένους ιστότοπους πρέπει να διατηρηθεί στην πλευρά του πελάτη ώστε να αποτραπεί η απώλεια της εμπιστευτικότητας των δεδομένων και να διατηρηθεί η ακεραιότητά τους. Η αναγκαιότητα της πολιτικής αυτής αναδεικνύεται στον παρακάτω παράδειγμα:

- Έστω ότι ο χρήστης X θέλει να ελέγξει τα email του στην υπηρεσία Gmail. Μετά την σύνδεση του στην υπηρεσία, επισκέπτεται ένα κακόβουλο ιστότοπο ανοίγοντας νέα καρτέλα στον περιηγητή. Η νέα σελίδα ελέγχει το πεδίο HTTP REFERER του HTTP αιτήματος και γνωρίζει ότι η προηγούμενη σελίδα ήταν το Gmail. Χρησιμοποιώντας cross-domain scripting, επιστρέφει πίσω στην σελίδα του Gmail και αποκτά πρόσβαση σε όλες τις επαφές του χρήστη και στο περιεχόμενο των email.

Για την αντιμετώπιση του παραπάνω περιορισμού και προκειμένου να επιτύχουμε την επικοινωνία του Geotyping και του εξωτερικού Geoparser απαιτείται η εγκατάσταση ενός τοπικού proxy server για τον Geocoder. Ο proxy δέχεται τα Ajax αιτήματα και τα προωθεί στον Geoparser.

6.1.3 Google Bulk Geocoding

Η υπηρεσία γεωκωδικοποίησης του Google Maps, θέτει συγκεκριμένους περιορισμούς για την συχνότητα και το πλήθος των αιτήσεων που μπορεί να δεχθεί από κάθε χρήστη. Συγκεκριμένα επιτρέπει μέχρι 2500 ερωτήματα την ημέρα για μεμονωμένους χρήστες από την ίδια διεύθυνση IP.

Κατά την ανάπτυξη της εφαρμογής, παρατηρήθηκε ότι η υπηρεσία επιτρέπει ένα μικρό αριθμό, λιγότερες από έξι, από ταυτόχρονες αιτήσεις γεωκωδικοποίησης. Στην

περίπτωση όμως που μια παράγραφος περιέχει περισσότερους όρους έπρεπε να επινοηθεί ένας μηχανισμός που θα επιτρέπει την γεωκωδικοποίηση όλων των όρων της παραγράφου.

Η λύση στο πρόβλημα ήρθε μέσα από την προσθήκη μιας ουράς αναμονής στον Geocoder του υποσυστήματος Google Maps. Κάθε όρος που πρέπει να γεωκωδικοποιηθεί εισάγεται στην ουρά. Στην συνέχεια, ο Geocoder εξάγει κάθε όρο από την ουρά στέλνοντας το αντίστοιχο αίτημα γεωκωδικοποίησης με συχνότητα 1 αίτημα / 0.5 sec. Τα αποτελέσματα προωθούνται στον χάρτη για την τοποθέτηση των δεικτών στις συντεταγμένες.

6.1.4 Ασάφεια κατά την γεωκωδικοποίηση

Η γεωκωδικοποίηση μια φράσης συχνά επιστρέφει μη αναμενόμενα αποτελέσματα. Το πρόβλημα αυτό οφείλεται στην ασάφεια των φράσεων και στα πολλαπλά νοήματα που μπορεί να έχουν. Χαρακτηριστικό παράδειγμα είναι η λέξη "Μοναστηράκι" που αποτελεί περιοχή της Στερεάς Ελλάδας, της Πελοποννήσου αλλά και κεντρική πλατεία της Αθήνας.

Μια quick-and-dirty λύση που επινοήσαμε, είναι ο περιορισμός της γεωκωδικοποίησης με την προσθήκη μιας φράσης που συγκεκριμενοποιεί την αρχική φράση που θέλουμε να εντοπίσουμε στον χάρτη. Για παράδειγμα αντί να στείλουμε την λέξη "Μοναστηράκι", στέλνουμε την φράση "Μοναστηράκι Αθήνα". Η φράση που προστίθεται ορίζεται μέσω της διεπαφής Geotyping Plug-in for CKEditor.

Η συγκεκριμένη τεχνική έχει ικανοποιητικά αποτελέσματα. Φυσικά ο χρήστης διατηρεί την δυνατότητα διόρθωσης των συντεταγμένων του κάθε δείκτη στον χάρτη. Μια καλύτερη λύση είναι πλέον εφικτή στην πρόσφατη έκδοση του Google Maps API η οποία επιτρέπει τον περιορισμό της γεωκωδικοποίησης μέσω της τεχνικής Viewport and Region Biasing. Περισσότερες λεπτομέρειες μπορείτε να βρείτε σε επόμενο κεφάλαιο και στην σελίδα του Google Maps API.

6.1.5 Υποστήριξη εικόνων

Το Geotyping έχει την δυνατότητα γεωκωδικοποίησης του κειμένου που συνοδεύει τους κόμβους των εικόνων σε ένα HTML έγγραφο. Ειδικότερα, όταν ο χρήστης εισάγει μια εικόνα, συμπληρώνει στο πεδίο "alt" του στοιχείου μια φράση που την περιγράφει. Η εφαρμογή εξάγει εκτός από το κείμενο της παραγράφου, το κείμενο από το πεδίο "alt" όλων των εικόνων της παραγράφου και το επεξεργάζεται με τον ίδιο ακριβώς τρόπο. Για κάθε όρο που εντοπίζεται στο πεδίο δημιουργείται ένα αόρατο στοιχείο που περιλαμβάνει τον όρο που έχει επισημανθεί.

6.1.6 Μαύρη λίστα

Ας υποθέσουμε το εξής σενάριο χρήσης. Ο συντάκτης πληκτρολογεί την φράση “Κατηφορίζουμε προς Αυστρία με τον E52, από τον οποίο βγαίνουμε για να πάρουμε τον 318 και μετά τον 30” . Ο Geoparser αναγνωρίζει σαν υποψήφιο γεωγραφικό όρο τα αναγνωριστικά E52 και 318. Στην συνέχεια ο Geocoder αντιστοιχίζει τους όρους αυτούς σε μη επιθυμητές συντεταγμένες. Ο χρήστης με την σειρά του διαγράφει τους δείκτες που δεν θέλει. Αργότερα επιστρέφει να διορθώσει την συγκεκριμένη φράση, η φράση δέχεται την προβλεπόμενη επεξεργασία και οι όροι αυτοί επανατοποθετούνται στην ίδια θέση. Επομένως ο χρήστης πρέπει να διαγράψει ξανά τους ανεπιθύμητους δείκτες.

Για να αντιμετωπίσουμε το πρόβλημα, ενσωματώσαμε στην εφαρμογή μια μαύρη λίστα στην οποία καταχωρούνται όροι που δεν θέλουμε να γεωκωδικοποιηθούν. Στην περίπτωση που εξετάσαμε, ο χρήστης εισάγει τους ανεπιθύμητους όρους στην μαύρη λίστα και εμποδίζει την περεταίρω γεωκωδικοποίησή τους.

6.2 Λεπτομέρειες υλοποίησης του συστήματος *Geotyping Text Management System*

6.2.1 *Registry*

Το συγκεκριμένο μοντέλο MVC χρησιμοποιεί ένα αντικείμενο registry στο οποίο αποθηκεύονται οι μεταβλητές που πρέπει να είναι προσβάσιμες από τον ελεγκτή που καλείται να εξυπηρετήσει το αίτημα. Πρόκειται ουσιαστικά για μια δομή που συγκεντρώνει και διαχωρίζει τις καθολικές και τις τοπικές μεταβλητές των κλάσεων που χρειάζεται ο ελεγκτής. Με αυτήν την τεχνική αποφεύγεται η υπερφόρτωση του καθολικού χώρου ονομάτων που αποτελεί συχνό φαινόμενο στα προγράμματα PHP.

6.2.2 *Διαβιβαστής και Search Engine Friendly Urls*

Όπως αναφέρθηκε σε προηγούμενο κεφάλαιο, το σύστημα έχει ως μοναδικό σημείο πρόσβασης το αρχείο `index.php`. Η πρόσβαση σε άλλα αρχεία και φακέλους απαγορεύεται με την κατάλληλη ρύθμιση του αρχείου `.htaccess` του διακομιστή. Κάθε αίτημα του χρήστη στέλνεται στο μοναδικό αυτό σημείο και αποτελείται ουσιαστικά από μια διεύθυνση της μορφής **`www.example.com/index.php?rt=articles/listing`**. Το αρχείο `index.php` μόλις δεχτεί ένα αίτημα, καλεί τον διαβιβαστή (router) δίνοντας του ως παράμετρο την συγκεκριμένη διεύθυνση. Ο router με την σειρά του, επιλέγει και εκκινεί τον κατάλληλο ελεγκτή βάσει της παραμέτρου `rt` της διεύθυνσης.

Ενισχύοντας ακόμα περισσότερο την παραπάνω λειτουργία, χρησιμοποιήσαμε το `mod_rewrite` του Apache Server ώστε διευθύνσεις της μορφής `www.example.com/articles/listing` να μεταφράζονται στην προηγούμενη μορφή. Η συγκεκριμένη μορφή διευθύνσεων πλεονεκτεί έναντι της πρώτης καθώς είναι φιλικότερη στον χρήστη και βοηθάει στην καλύτερη κατάταξη της σελίδας στις μηχανές αναζήτησης αφού περιέχει σημασιολογική πληροφορία για το περιεχόμενο της κάθε σελίδας.

6.3 Πλατφόρμες και προγραμματιστικά εργαλεία

Στην παρούσα ενότητα αναφέρονται τα προγραμματιστικά εργαλεία και οι πλατφόρμες ανάπτυξης λογισμικού που χρησιμοποιήθηκαν κατά την φάση της ανάπτυξης των συστημάτων της διπλωματικής εργασίας. Η εφαρμογή που αναπτύχθηκε είναι μία δικτυακή εφαρμογή, συνεπώς τα εργαλεία που παρουσιάζονται περιστρέφονται γύρω από την ανάπτυξη και παροχή υπηρεσιών στον διαδίκτυο.

6.3.1 Apache HTTP Server

Ο Apache HTTP Server είναι ένας διακομιστής ιστοσελίδων ο οποίος διαδραμάτισε προεξέχοντα ρόλο στην αρχική ανάπτυξη του παγκόσμιου ιστού. Μέχρι σήμερα, είναι ο πιο διαδεδομένος διακομιστής με περισσότερα από 100 εκατομμύρια ιστοσελίδες να βασίζονται τη λειτουργία τους σε αυτόν. Τα χαρακτηριστικά που διαθέτει του επιτρέπουν να ανταγωνίζεται τους υπόλοιπους διαθέσιμους διακομιστές τόσο σε όρους παρεχόμενων λειτουργιών όσο και σε απόδοση. Ο Apache διατίθεται για μία μεγάλη πληθώρα λειτουργικών συστημάτων, όπως Windows, Linux, FreeBSD, κλπ.. Παρόλα αυτά στις περισσότερες των περιπτώσεων προτιμάται η λειτουργία του σε συνδυασμό με το λειτουργικό σύστημα Linux. Αποτελεί ανοικτό λογισμικό υπό την άδεια Apache License και η ανάπτυξη και διαχείρισή του πραγματοποιείται από μία ανοικτή κοινότητα προγραμματιστών υπό την αιγίδα του Apache Software Foundation. Ο Apache HTTP Server υποστηρίζει ένα ευρύ σύνολο χαρακτηριστικών, πολλά από τα οποία υλοποιούνται ως ξεχωριστά modules τα οποία και επεκτείνουν το βασικό σύνολο λειτουργιών του διακομιστή. Ενδεικτικά αναφέρονται οι εξής δυνατότητες:

- Υποστήριξη γλωσσών προγραμματισμού στην πλευρά του διακομιστή, όπως είναι PHP, Perl, Python, κλπ.
- Δυνατότητα ταυτοποίησης χρηστών μέσω διάφορων διαδεδομένων συστημάτων ταυτοποίησης.
- Τροποποίηση των διευθύνσεων ώστε να είναι φιλικότερες στον χρήστη και στις μηχανές αναζήτησης.
- Συμπίεση των αποστέλλομενων δεδομένων για βελτίωση των επιδόσεων του διακομιστή.
- Υποστήριξη ασφαλούς μεταφοράς της πληροφορίας μέσω του πρωτοκόλλου TLS.

Αν και ο βασικός σχεδιαστικός στόχος του Apache δεν ήταν η επίτευξη των μέγιστων επιδόσεων, παρόλα αυτά εξακολουθεί να είναι ένας από τους πιο γρήγορους διακομιστές που κυκλοφορούν σήμερα. Έχει τη δυνατότητα μέσω του αντίστοιχου module να υποστηρίζει

διαφορετικές αρχιτεκτονικές υπολογιστικών συστημάτων με αποτέλεσμα να εκμεταλλεύεται κατά το μέγιστο δυνατό τρόπο την εκάστοτε υπολογιστική υποδομή.

Έχοντας υπόψη το σύνολο των παραπάνω αναφερόμενων πληροφοριών για τον Apache HTTP Server πραγματοποιήθηκε η επιλογή του ως διακομιστή για τη δικτυακή εφαρμογή Geotyping Text Management System που αναπτύχθηκε στα πλαίσια της παρούσας διπλωματικής εργασίας.

6.3.2 *MySQL*

Η MySQL είναι ένα διαδεδωμένο σύστημα σχεσιακής βάσης δεδομένων που αριθμεί σήμερα περισσότερες από 6 εκατομμύρια εγκαταστάσεις παγκοσμίως. Λειτουργεί ως server και, επομένως, μπορεί να εξυπηρετήσει την πρόσβαση πολλαπλών χρηστών σε πολλαπλές βάσεις δεδομένων ταυτόχρονα. Είναι ανοικτό λογισμικό μιας και διανέμεται υπό την άδεια GNU General Public License. Ανήκει στην κερδοσκοπική εταιρία MySQL AB, η οποία πλέον είναι θυγατρική εταιρία της Sun Microsystems, η οποία με τη σειρά της βρίσκεται στη διαδικασία απορρόφησης από την Oracle Corporation. Η MySQL χρησιμοποιείται πολύ συχνά σε έργα ανοικτού λογισμικού και, συνήθως, συνδυάζεται με το λειτουργικό σύστημα Linux και το διακομιστή ιστοσελίδων Apache HTTP Server. Ένας σημαντικός αριθμός από μεγάλης κλίμακας διαδικτυακών προϊόντων χρησιμοποιεί τη MySQL, όπως Wikipedia, Google και Facebook. Η MySQL είναι σχεδιασμένη ώστε να υποστηρίζει ένα μεγάλο αριθμό από λειτουργικά συστήματα περιλαμβανομένων των Linux, Windows, NetBSD, OpenBSD, κλπ.. Επιπλέον, διατίθενται βιβλιοθήκες πρόσβασης στο σύστημα της βάσης δεδομένων για τις περισσότερες γνωστές γλώσσες προγραμματισμού. Τα χαρακτηριστικά που διαθέτει καλύπτουν ένα ευρύ φάσμα εφαρμογών. Ενδεικτικά παρατίθενται τα εξής:

- Υποστήριξη του ANSI SQL 99.
- Υποστήριξη για stored procedures, cursors και triggers.
- Υποστήριξη για updatable views.
- Παρέχει πολλαπλές ανεξάρτητες μηχανές αποθήκευσης δεδομένων, όπως MyISAM, InnoDB, κλπ.. Η κάθε μηχανή είναι βελτιστοποιημένη για συγκεκριμένη χρήση.
- Υποστήριξη Full-text indexing.

Με βάση τα παραπάνω, επιλέχθηκε να χρησιμοποιηθεί η MySQL ως το σύστημα σχεσιακής βάσης δεδομένων το οποίο ανέλαβε το ρόλο της διαχείρισης των δεδομένων της εφαρμογής διαχείρισης κειμένου.

6.3.3 PHP

Η PHP είναι μία ευρέως διαδεδομένη scripting γλώσσα προγραμματισμού, η οποία ειδικεύεται και χρησιμοποιείται κατά κόρον στην ανάπτυξη εφαρμογών για τον παγκόσμιο ιστό. Είναι συμβατή με μία ευρεία συλλογή από λειτουργικά συστήματα και διακομιστές ιστοσελίδων με αποτέλεσμα να χρησιμοποιείται σε περισσότερες από 20 εκατομμύρια ιστοσελίδες παγκοσμίως. Αποτελεί ανοικτό λογισμικό και διανέμεται υπό την άδεια PHP License. Αξίζει να σημειωθεί ότι, αν και ο αρχικός στόχος της γλώσσας ήταν αποκλειστικά η υποστήριξη δημιουργίας δυναμικών ιστοσελίδων, σήμερα μπορεί να χρησιμοποιηθεί για τη συγγραφή ανεξάρτητων εφαρμογών είτε γραμμής εντολών είτε γραφικού interface.

Από την έκδοση 4 και έπειτα, τα αρχεία πηγαίου κώδικα μεταγλωττίζονται αυτόματα σε bytecode και εκτελούνται από την αφηρημένη μηχανή Zend Engine με αποτέλεσμα αυξημένες επιδόσεις σε σχέσεις με τις προγενέστερες εκδόσεις. Παράλληλα, υπάρχουν λύσεις σε επίπεδο λογισμικού που αξιοποιούν ένα σύνολο από τεχνικές για να επιτύχουν ακόμη μεγαλύτερες επιδόσεις όταν αυτό κρίνεται αναγκαίο για την εύρυθμη εξυπηρέτηση του συνόλου των χρηστών της δικτυακής εφαρμογής. Άλλο ένα σημαντικό χαρακτηριστικό της γλώσσας προγραμματισμού PHP είναι ότι από την έκδοση 5 και έπειτα υποστηρίζει πλήρως το αντικειμενοστραφές μοντέλο προγραμματισμού με αποτέλεσμα ο προγραμματιστής να έχει την ευχέρεια να αξιοποιήσει σύγχρονες τεχνικές ανάπτυξης λογισμικού για την συγγραφή της εκάστοτε εφαρμογής.

Στη βάση όλων των παραπάνω χαρακτηριστικών, η PHP επιλέχθηκε ως η γλώσσα ανάπτυξης του συστήματος διαχείρισης κειμένου της παρούσας διπλωματικής εργασίας. Επιπλέον, κρίθηκε ως θετικό στοιχείο η δυνατότητα άριστης συνεργασίας τόσο με τον Apache HTTP Server όσο και με το σύστημα σχεσιακής βάσης δεδομένων MySQL.

6.3.4 JavaScript, Βιβλιοθήκες και APIs

Όπως αναφέρθηκε παραπάνω, η εφαρμογή Geotyping και μέρος του συστήματος Geotyping Text Management System είναι σε γλώσσα JavaScript. Ειδικότερα η εφαρμογή Geotyping στηρίχθηκε στην δημοφιλή βιβλιοθήκη jQuery, στο Google Maps API και στο CKEditor API. Η εξασφάλιση της αρμονικής συνεργασίας των παραπάνω βιβλιοθηκών οδήγησε στην υλοποίηση των απαιτούμενων λειτουργιών της εφαρμογής. Παράλληλα προσπαθήσαμε να αξιοποιήσουμε με βέλτιστο τρόπο τα πλεονεκτήματα της κάθε μία βιβλιοθήκης στα διαφορετικά υποσυστήματα της εφαρμογής.



Ολοκληρώνοντας την περιγραφή των βιβλιοθηκών και των εργαλείων της JavaScript αξίζει να αναφερθούμε και στο Fancybox, ένα εργαλείο για την απεικόνιση περιεχόμενου σε αναδυόμενο παράθυρα βασισμένο σε jQuery.

6.3.5 Πλατφόρμα ανάπτυξης

Η ανάπτυξη της εφαρμογής Geotyping έγινε στην πλατφόρμα Eclipse, έκδοση Galileo με εγκατεστημένο το Aptana Plug-in που παρέχει χρήσιμα εργαλεία αποσφαλμάτωσης για την JavaScript.



Το ολοκληρωμένο σύστημα διαχείρισης κειμένου Geotyping Text Management System αναπτύχθηκε στην πλατφόρμα Netbeans 6.9.1 λόγω των πλεονεκτημάτων και των εργαλείων που προσφέρει στην χρήση της PHP.



7. Έλεγχος και Εγκατάσταση

Στο παρόν κεφάλαιο παρουσιάζεται η εγκατάσταση του συστήματος και οι λειτουργίες της τελικής εφαρμογής. Το κεφάλαιο αυτό, λειτουργεί ως εγχειρίδιο .

7.1 Εγκατάσταση

7.1.1 Απαιτήσεις

- Apache 2.2 με ενεργοποιημένο το mod_rewrite module
- MySQL 5.5 η νεώτερη
- PHP Version 5.3 η νεώτερη με εγκατεστημένη την βιβλιοθήκη cURL
- Προαιρετικά εγκαταστήστε έναν mail server στο μηχάνημά σας.

Για το λειτουργικό σύστημα Windows προτείνεται η χρήση του πακέτου **XAMPP** που περιλαμβάνει όλα τα παραπάνω σε ένα πακέτο. Το XAMPP μπορείτε να το κατεβάσετε από την διεύθυνση <http://www.apachefriends.org/en/xampp.html>

7.1.2 Οδηγίες εγκατάστασης

Βήμα 1: Εγκατάσταση βάσης δεδομένων

Ανοίξτε με τον editor της επιλογής σας το αρχείο **models/db.class.php** και τροποποιήστε τις τιμές των παρακάτω μεταβλητών ανάλογα με τις ρυθμίσεις του διακομιστή της βάσης δεδομένων που θα φιλοξενήσει την εφαρμογή.

```
$db_host="localhost";
```

```
$db_user = "root";
```

```
$db_pass = "";
```

```
$db_name = "gtms";
```

```
$db_port = "";
```

```
$db_table_prefix = "";
```

Βήμα 2: Ρυθμίσεις εφαρμογής

Τροποποιήστε το αρχείο **includes/config.php** αλλάζοντας τις καθολικές σταθερές ανάλογα με τις ρυθμίσεις του τοπικού Apache Server.

Βήμα 3: Εγκατάσταση εφαρμογής

Ανοίξτε έναν browser και μεταβείτε στην διεύθυνση *Your_server_address/Path_to_app/install/*

Βήμα 4: Ολοκλήρωση εγκατάστασης

Ακολουθείστε τις οδηγίες της σελίδας και μόλις ολοκληρωθεί η εγκατάσταση διαγράψτε τον φάκελο **install**.

7.1.3 Εγκατάσταση Geotyping

Η εφαρμογή Geotyping μπορεί να ενσωματωθεί σε οποιοδήποτε σύστημα διαχείρισης περιεχομένου. Διανέμεται σε zip μορφή και η εγκατάσταση γίνεται με την φόρτωση των αρχείων JavaScript της εφαρμογής όπως φαίνεται στο index.html που περιλαμβάνεται στο πακέτο.

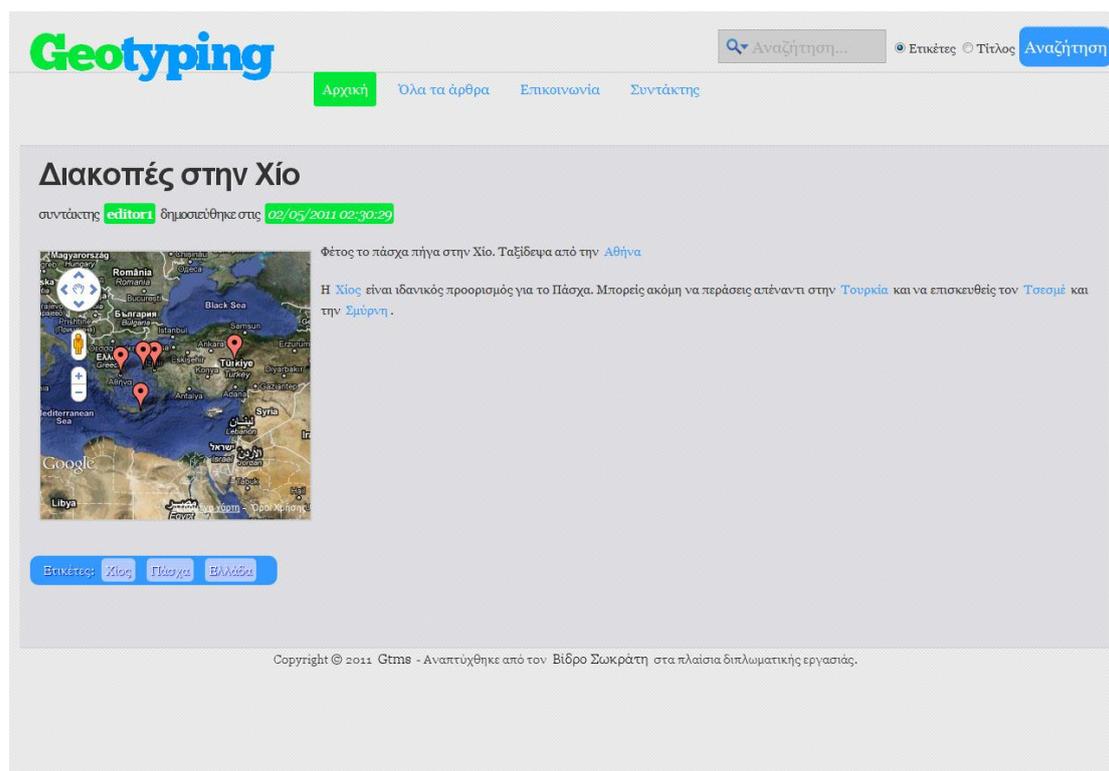
7.2 Browser Compatibility

Η εφαρμογή αναπτύχθηκε και δοκιμάστηκε στο περιηγητή Google Chrome 11. Έχει ελεγχθεί η συμβατότητα της με τον περιηγητή Mozilla Firefox 4. Ένα ιδιαίτερα χρήσιμο εργαλείο για την αποσφαλμάτωση (debugging) και τον έλεγχο ορθής λειτουργίας της εφαρμογής ήταν η κονσόλα του Google Chrome και η επέκταση FireBug του περιηγητή Mozilla Firefox 4.

7.3 Εκκίνηση και χρήση της εφαρμογής

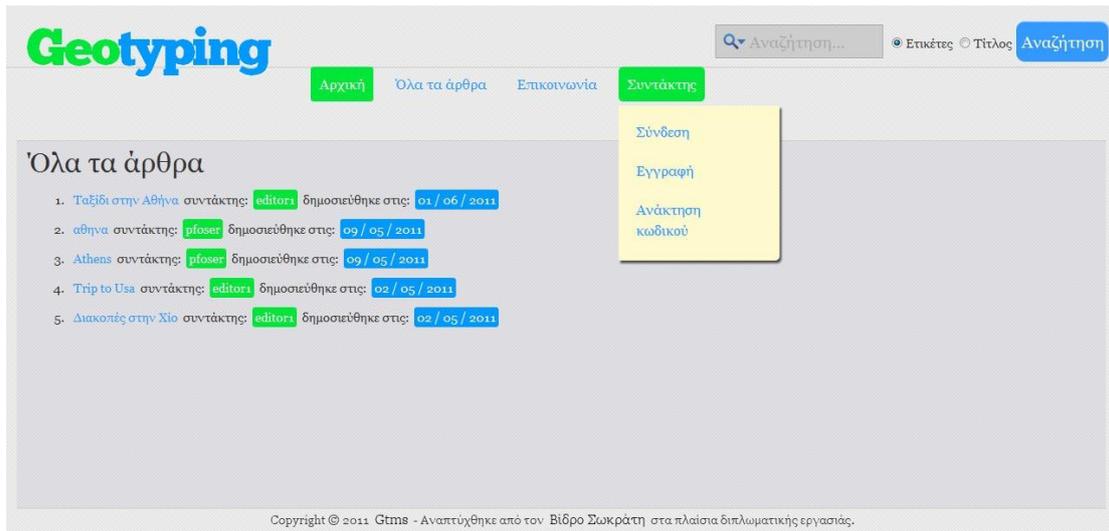
Στο τρέχων κεφάλαιο θα παρουσιάσουμε την χρήση του συστήματος διαχείρισης κειμένου Geotyping Text Management System που περιλαμβάνει την εφαρμογή Geotyping.

Εκκινώντας το σύστημα, παρατηρούμε τα στοιχεία αλληλεπίδρασης που συνθέτουν το γραφικό περιβάλλον του συστήματος διαχείρισης κειμένου. Ο χρήστης έχει την δυνατότητα πλοήγησης, μέσω του βασικού μενού που βρίσκεται στο πάνω μέρος της σελίδας και αναζήτησης άρθρων μέσω της φόρμας αναζήτησης. Τα παραπάνω αποτελούν τα σταθερά στοιχεία του γραφικού περιβάλλοντος και περιλαμβάνονται σε κάθε σελίδα. Στο κεντρικό τμήμα της εφαρμογής προβάλλονται τα περιεχόμενα της κάθε σελίδας που επιλέγει ο χρήστης. Η αρχική σελίδα προβάλλει ένα τυχαίο άρθρο (εικόνα 7.1). Παρατηρούμε ότι στο άρθρο και στον χάρτη που το συνοδεύει έχουν επισημανθεί η γεωγραφικοί όροι που εντοπίστηκαν στο κείμενο. Στις επόμενες δύο εικόνες προβάλλεται η λίστα όλων των δημοσιευμένων άρθρων (εικόνα 7.2) και η φόρμα επικοινωνίας (εικόνα 7.3).

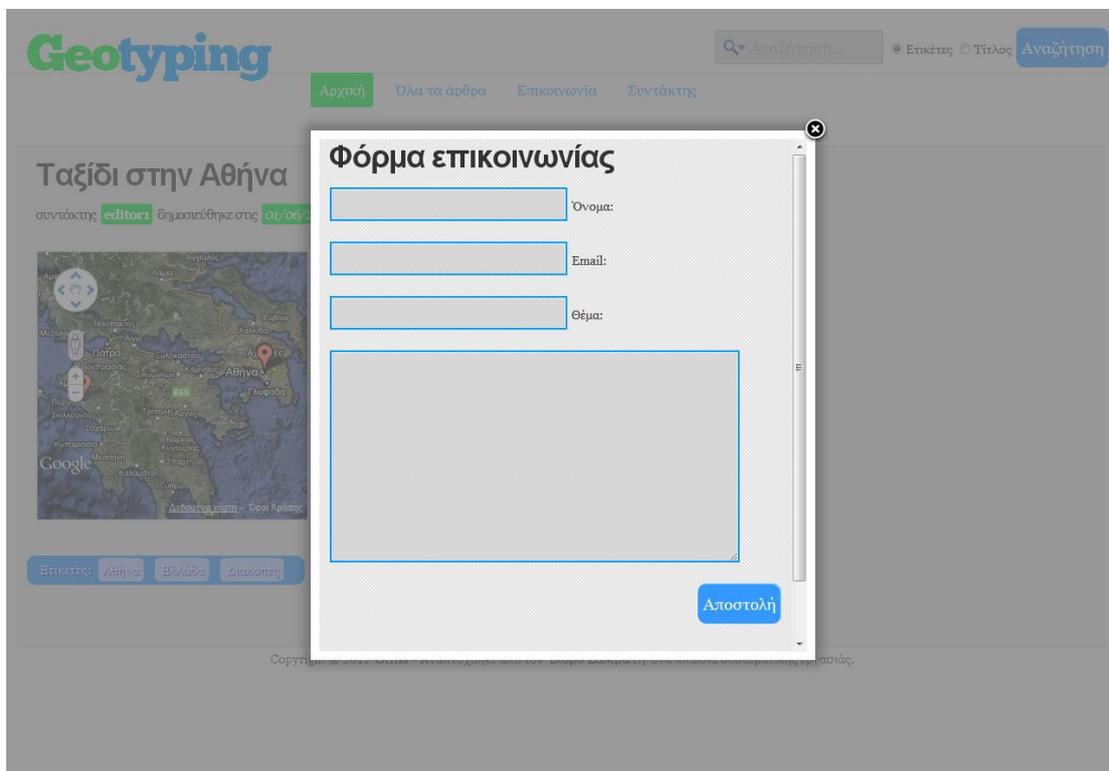


The screenshot displays the Geotyping Text Management System interface. At the top, the 'Geotyping' logo is on the left, and a search bar with the text 'Αναζήτηση...' is on the right. Below the logo, a navigation menu includes 'Αρχική', 'Όλα τα άρθρα', 'Επικοινωνία', and 'Συντάκτης'. The main content area features an article titled 'Διακοπές στην Χίο' with a sub-header 'συντάκτης editor1 δημοσιεύθηκε στις 02/05/2011 02:30:29'. The article text includes a map of the Aegean region with red pins marking locations like 'Αθήνα', 'Χίος', 'Τουρκία', and 'Τσομό'. Below the map, there are tags for 'Χίος', 'Πάσχα', and 'Βιλλιά'. At the bottom, a copyright notice reads 'Copyright © 2011 Gtms - Αναπτύχθηκε από τον Βίδρο Σωκράτη στα πλαίσια διπλωματικής εργασίας.'

7-1 Αρχική σελίδα Geotyping Text Management System



7-2 Λίστα δημοσιευμένων άρθρων



7-3 Φόρμα επικοινωνίας

Η εγγραφή ενός χρήστη-συντάκτη στο σύστημα γίνεται μέσω της φόρμας εγγραφής (εικόνα 7.4). Τα υποχρεωτικά πεδία που πρέπει να συμπληρωθούν είναι το όνομα χρήστη, ο κωδικός και το email. Μετά από μια επιτυχημένη εγγραφή, ο χρήστης μπορεί πλέον να συνδεθεί στο σύστημα (εικόνα 7.6). Σε περίπτωση που ο χρήστης ξεχάσει τον κωδικό πρόσβασης, μπορεί να ανακτήσει το συνηματικό του μέσω της φόρμας ανάκτησης κωδικού (εικόνα 7.5).

Geotyping [Ετικέτες](#) [Τίτλος](#) [Αναζήτηση](#)

[Αρχική](#) [Όλα τα άρθρα](#) [Επικοινωνία](#) [Συντάκτης](#)

Εγγραφή συντάκτη

Όνομα χρήστη:

Κωδικός:

Επιβεβαίωση:

Email:

Copyright © 2011 Gtms - Αναπτύχθηκε από τον Βίδρο Σωκράτη στα πλαίσια διπλωματικής εργασίας.

7-4 Σελίδα εγγραφής συντάκτη

Geotyping [Ετικέτες](#) [Τίτλος](#) [Αναζήτηση](#)

[Αρχική](#) [Όλα τα άρθρα](#) [Επικοινωνία](#) [Συντάκτης](#)

Ανάκτηση κωδικού

Όνομα χρήστη:

Email:

Copyright © 2011 Gtms - Αναπτύχθηκε από τον Βίδρο Σωκράτη στα πλαίσια διπλωματικής εργασίας.

7-5 Σελίδα ανάκτησης κωδικού

Geotyping [Ετικέτες](#) [Τίτλος](#) [Αναζήτηση](#)

[Αρχική](#) [Όλα τα άρθρα](#) [Επικοινωνία](#) [Συντάκτης](#)

Σύνδεση συντάκτη

Όνομα χρήστη:

Κωδικός:

Copyright © 2011 Gtms - Αναπτύχθηκε από τον Βίδρο Σωκράτη στα πλαίσια διπλωματικής εργασίας.

7-6 Σελίδα σύνδεσης συντάκτη

Στα screenshots που ακολουθούν, παρουσιάζονται όλες οι σελίδες στις οποίες ο χρήστης αποκτά πρόσβαση μετά την επιτυχή σύνδεση του. Στην πρώτη σελίδα, εμφανίζεται η λίστα με όλες τις δημοσιεύσεις του συνδεδεμένου χρήστη (εικόνα 7.7). Ο χρήστης από αυτό το σημείο μπορεί να δημιουργήσει ένα νέο άρθρο, να επεξεργαστεί ή να διαγράψει ένα παλαιότερο και να διαχειριστεί τα στοιχεία του λογαριασμού του (εικόνα 7.8).

The screenshot shows the Geotyping website interface. At the top, there is a search bar with the text "Αναζήτηση..." and a blue "Αναζήτηση" button. Below the search bar, there are navigation tabs: "Αρχική" (highlighted in green), "Όλα τα άρθρα", "Επικοινωνία", and "Συντάκτης". The main content area displays a greeting "Καλώς όρισες!" followed by "Τα άρθρα μου". Below this is a table with the following data:

Τίτλος	Ημερομηνία δημοσίευσης	Δημοσιευμένο	Ενέργειες
Ταξίδι στην Αθήνα	01/06/2011 02:14:14 AM	✓	Επεξεργασία Διαγραφή
Trip to Usa	02/05/2011 02:39:39 AM	✓	Επεξεργασία Διαγραφή
Διακοπές στην Χίο	02/05/2011 02:30:29 AM	✓	Επεξεργασία Διαγραφή

At the bottom right of the table area, there is a link "Νέα δημοσίευση". At the very bottom of the page, there is a copyright notice: "Copyright © 2011 Gtms - Αναπτύχθηκε από τον Βίδρο Σωκράτη στα πλαίσια διπλωματικής εργασίας."

7-7 Αρχική σελίδα συντάκτη

The screenshot shows the Geotyping website interface for account settings. At the top, there is a search bar with the text "Αναζήτηση..." and a blue "Αναζήτηση" button. Below the search bar, there are navigation tabs: "Αρχική", "Όλα τα άρθρα", "Επικοινωνία", and "Συντάκτης". The main content area displays the title "Ο λογαριασμός μου" and the following information:

Το όνομα μου είναι: **editor1**
 Το όνομα της ομάδας μου είναι: **Editor**
 Η εγγραφή έγινε στις **02-05-2011**
 Η τελευταία φορά που συνδέθηκα ήταν στις **05-06-2011 21:54**

Below this information, there is a section titled "Αλλαγή διεύθυνσης email". It includes a label "Email:" and a text input field containing "editor1@gtms.gr". Below the input field is a blue button labeled "Αλλαγή Email".

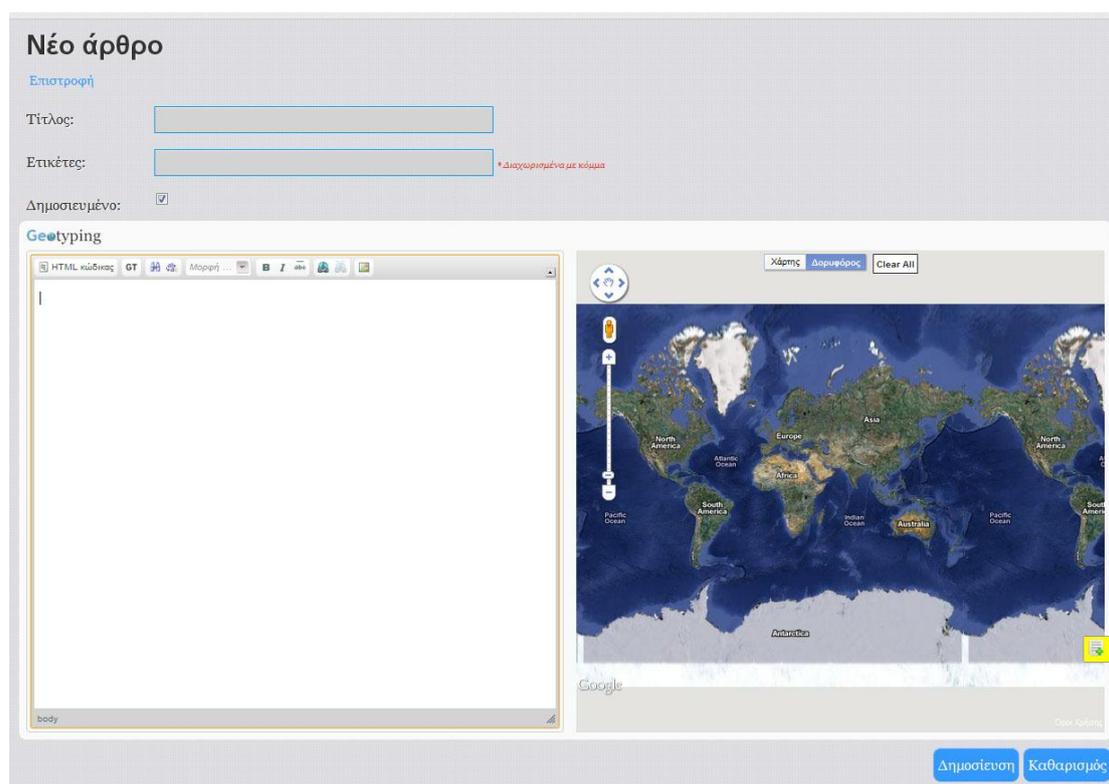
Below the email section, there is a section titled "Αλλαγή κωδικού". It includes three text input fields: "Τρέχον κωδικός:", "Νέος κωδικός:", and "Επιβεβαίωση:". Below the input fields is a blue button labeled "Αλλαγή κωδικού".

7-8 Ρυθμίσεις λογαριασμού

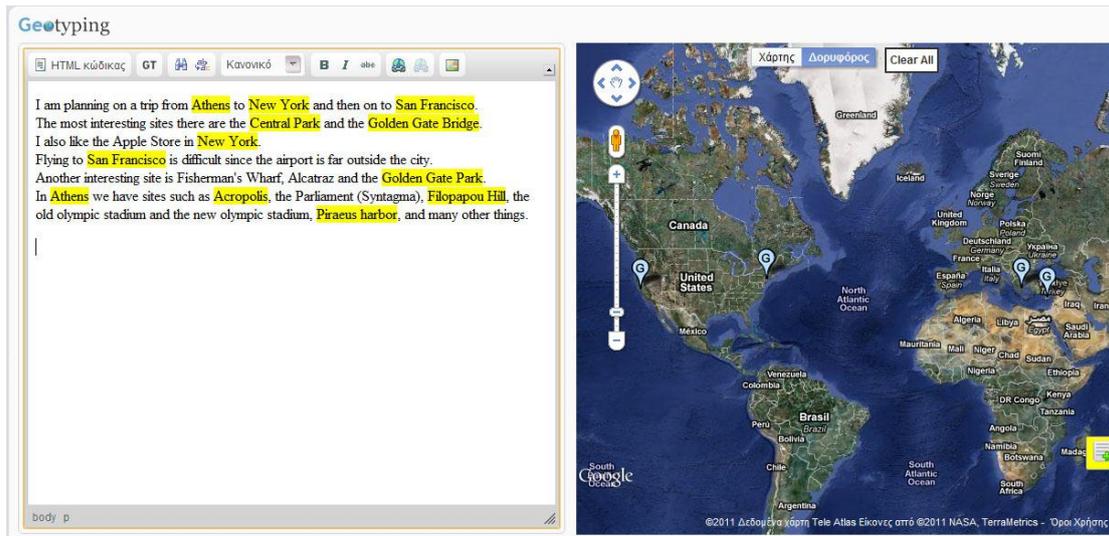
Η σελίδα δημιουργία νέου άρθρου αποτελεί την καρδιά του συστήματος (εικόνα 7.9). Στην σελίδα έχει ενσωματωθεί η εφαρμογή Geotyping. Αρχικά ο χρήστης καλείται να συμπληρώσει τον τίτλο του άρθρου και τις σχετικές με το περιεχόμενο ετικέτες. Στην συνέχεια εισάγει το περιεχόμενο του άρθρου μέσω του Geotyping.

Στις επόμενες παραγράφους περιγράφουμε αναλυτικά τον τρόπο λειτουργίας της εφαρμογής Geotyping. Παρατηρούμε ότι η εφαρμογή αποτελείται από τον επεξεργαστή κειμένου στα αριστερά και από τον γεωγραφικό χάρτη στα δεξιά (εικόνα 7.9). Ένα στιγμιότυπο της εφαρμογής έπειτα από την ολοκλήρωση της συγγραφής ενός άρθρου δίνεται στην εικόνα 7.10. Στο κείμενο έχουν επισημανθεί οι γεωγραφικοί όροι που εντοπίστηκαν και έχουν τοποθετηθεί οι αντίστοιχοι δείκτες στον χάρτη. Ο χρήστης μπορεί να έχει μια εποπτική εικόνα για όλους τους δείκτες του χάρτη μέσω της λίστας δεικτών που εμφανίζεται στα δεξιά του χάρτη (εικόνα 7.11).

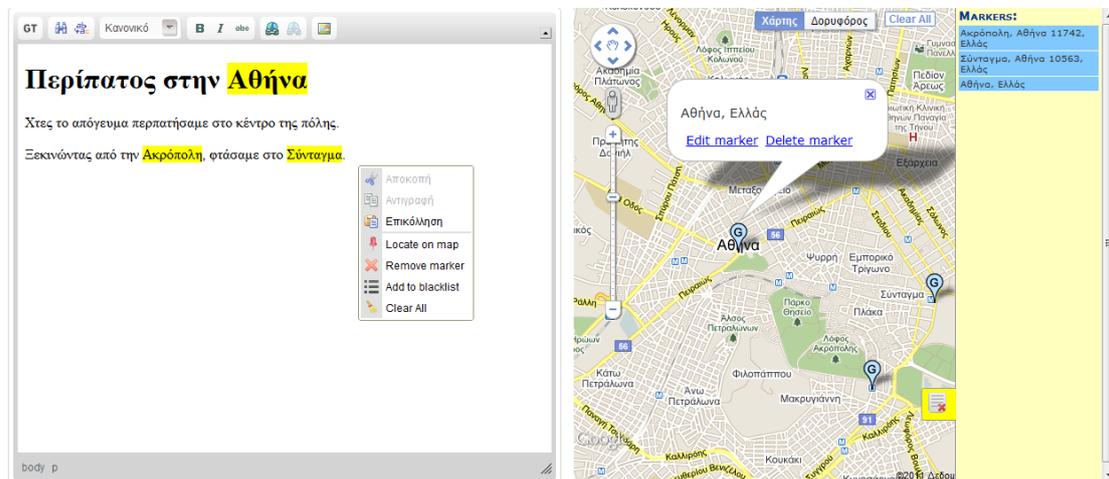
Η γεωκωδικοποίηση κειμένου γίνεται αυτόματα από την εφαρμογή. Το μενού στον περιοχή κειμένου (εικόνα 7.11) αποτελεί την διεπαφή του χρήστη για την λειτουργία της χειροκίνητης γεωκωδικοποίησης και της επεξεργασίας όρων και δεικτών στον χάρτη.



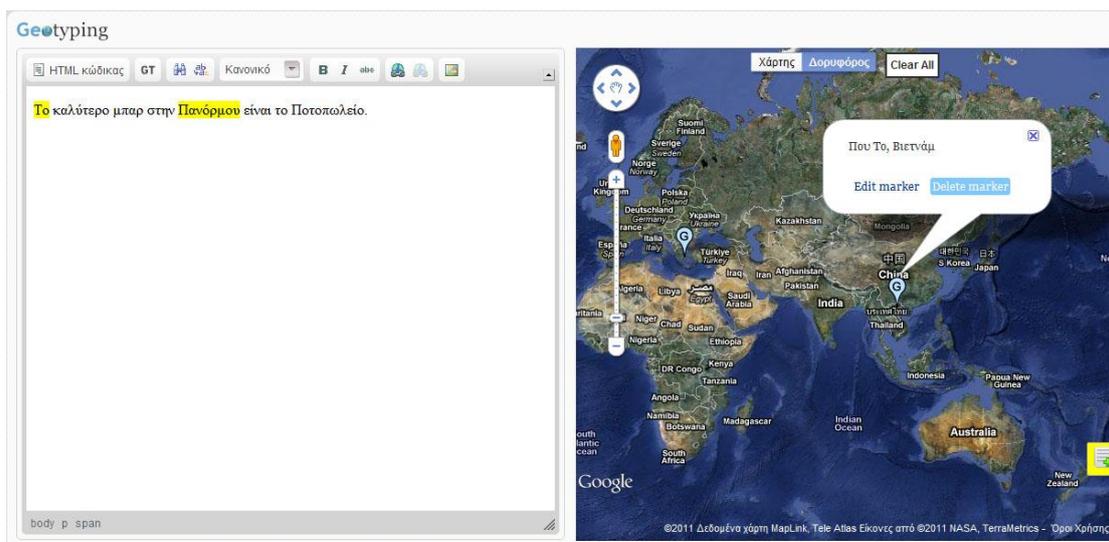
7-9 Δημοσίευση νέου άρθρου



7-10 Γραφικό περιβάλλον του Geotyping



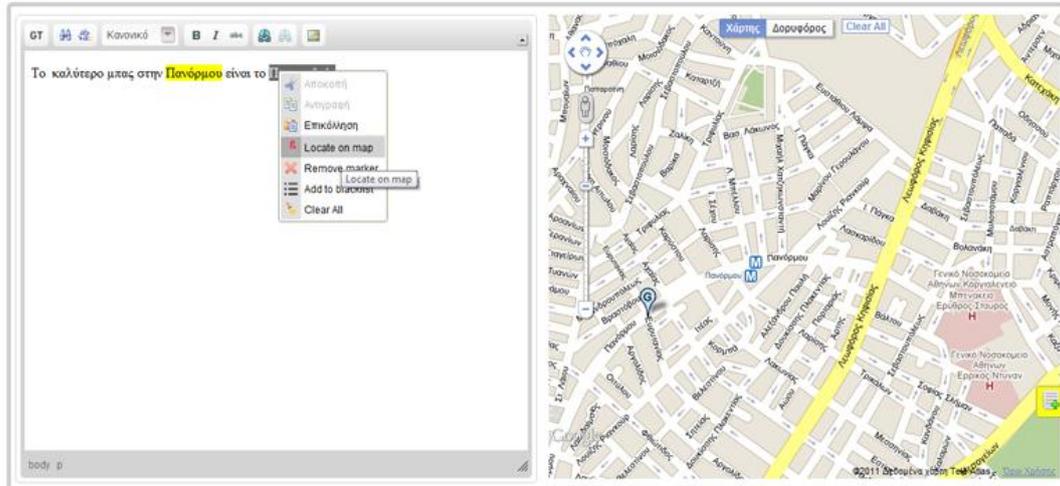
7-11 Τα μενού της εφαρμογής



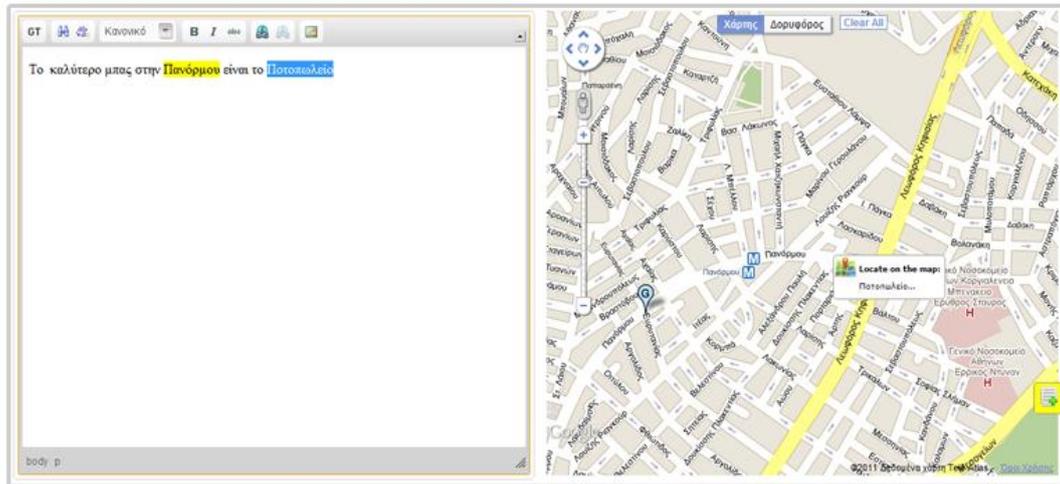
7-12 Διαγραφή όρου

Στην επόμενη εικόνα (εικόνα 7.13) παρουσιάζεται η λειτουργία χειροκίνητης απεικόνισης που προσφέρει το Geotyping. Υποθέτουμε ότι ο χρήστης θέλει να αντιστοιχίσει σε συντεταγμένες στον χάρτη την λέξη «Ποτοπωλείο» που είναι μια συγκεκριμένη καφετέρια την οποία δεν γνωρίζει ο Geocoder. Ο χρήστης μαρκάρει την λέξη μέσω του ποντικιού, κάνει δεξί κλικ και επιλέγει από το εμφανιζόμενο μενού την επιλογή «Locate on map». Στην συνέχεια, εάν η επιλογή είναι έγκυρη, εμφανίζεται ένα πλαίσιο που περιέχει την επιλεγμένη φράση. Το πλαίσιο αυτό πρέπει να τοποθετηθεί στην επιθυμητή θέση στον χάρτη. Τα βήματα της διαδικασίας παρουσιάζονται αναλυτικά.

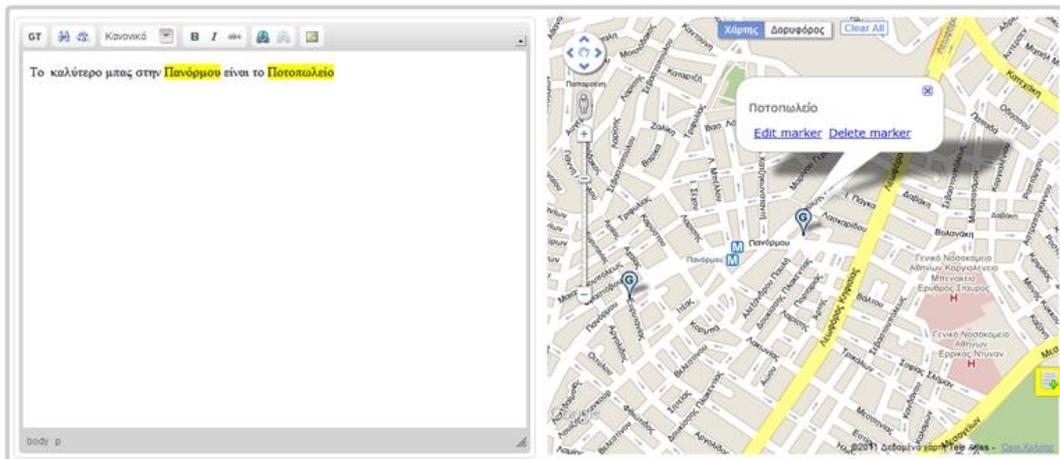
Βήμα 1



Βήμα 2



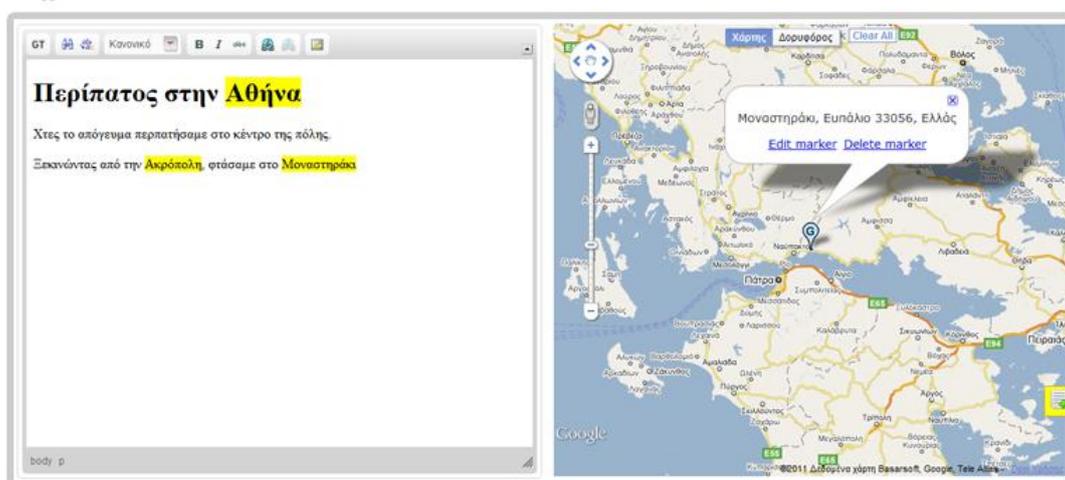
Βήμα 3



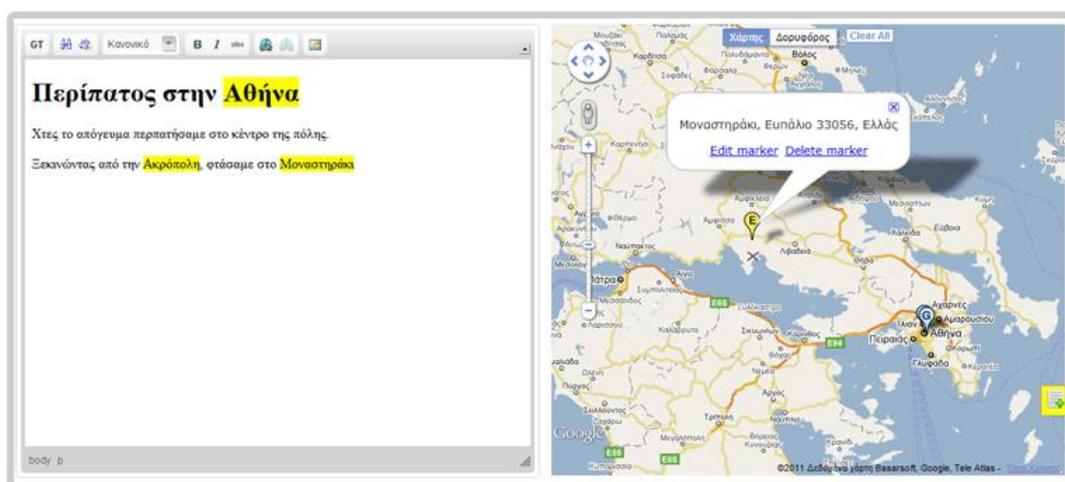
7-13 Χειροκίνητη γεωκωδικοποίηση φράσης

Ο Geocoder αρκετές φορές επιστρέφει μη αναμενόμενα αποτελέσματα λόγω της ασάφειας και των πολλαπλών νοημάτων που πιθανός να έχουν οι λέξεις και τα τοπωνύμια σε μία γλώσσα. Ένα χαρακτηριστικό παράδειγμα είναι το ακόλουθο (εικόνα 7.14). Ο χρήστης πληκτρολογεί την φράση «Χτες το απόγευμα περπατήσαμε στο κέντρο της Αθήνας. Ξεκινώντας από την Ακρόπολη καταλήξαμε στο Μοναστηράκι.» Η λέξη Μοναστηράκι εκτός από κεντρική πλατεία της Αθήνας, είναι και περιοχή στην Στερεά Ελλάδα, συνεπώς ο Geocoder ενδέχεται να επιστρέψει μη αναμενόμενο αποτέλεσμα. Ο χρήστης εντοπίζει το σφάλμα και διορθώνει τον δείκτη τοποθετώντας τον στο σημείο που επιθυμεί.

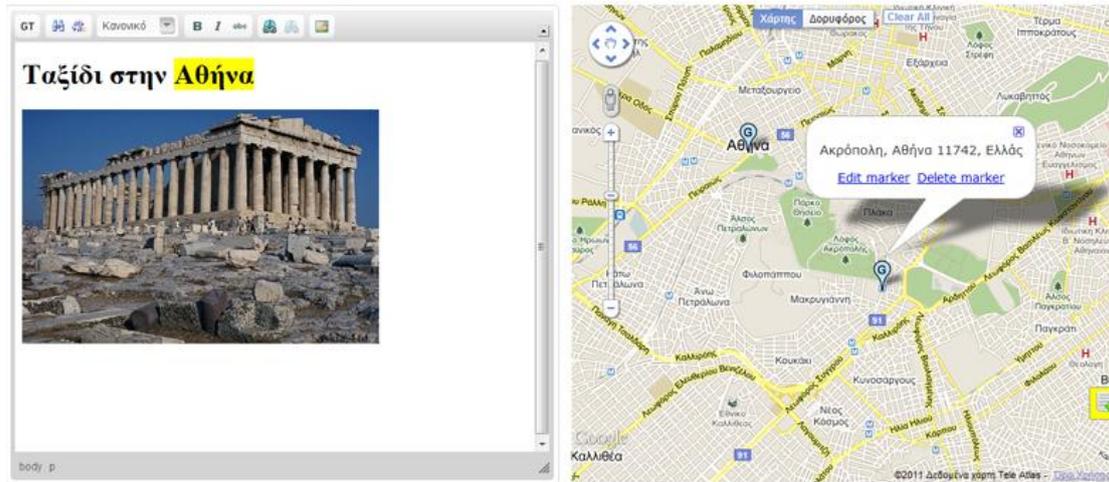
Βήμα 1



Βήμα 2

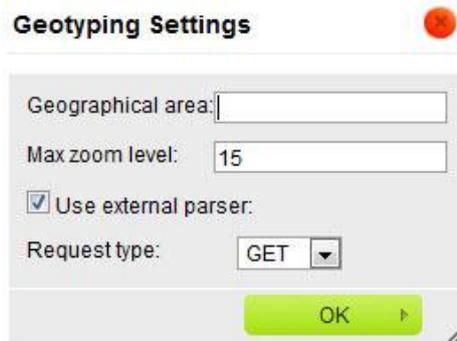


7-14 Διόρθωση γεωγραφικού όρου στον χάρτη



7-15 Υποστήριξη μορφοποιημένου κειμένου και εικόνων

Ο χρήστης μπορεί να ρυθμίσει τις βασικές παραμέτρους λειτουργίας του Geotyping μέσω του αναδυόμενου παραθύρου του υποσυστήματος Geotyping Plug-in for CKEditor (εικόνα 7.15). Οι παράμετροι που περιλαμβάνονται είναι η ευρύτερη γεωγραφική περιοχή που σχετίζεται με το κείμενο, το μέγιστο επίπεδο εστίασης στον χάρτη, η χρήση ή όχι του εξωτερικού Geoparser καθώς και ο τύπος των αιτημάτων προς αυτόν.



7-16 Geotyping Plug-in for CKEditor

8. *Επίλογος*

Στο τελευταίο κεφάλαιο γίνεται μια ανακεφαλαίωση της διπλωματικής εργασίας. Συνοψίζονται οι στόχοι που τέθηκαν και οι τεχνικές που επιστρατεύτηκαν ώστε να καλυφθούν οι απαιτήσεις του συστήματος. Επιπλέον, τεκμηριώνεται μία σειρά από συμπεράσματα τα οποία προέκυψαν μέσα από τη διαδικασία της ανάπτυξης του λογισμικού της διπλωματικής εργασίας. Τέλος, προτείνονται μελλοντικές επεκτάσεις και βελτιώσεις του συστήματος που στοχεύουν στην επέκταση και στην βελτιστοποίηση των ήδη υποστηριζόμενων λειτουργιών της εφαρμογής Geotyping.

8.1 *Σύνοψη και συμπεράσματα*

Αντικείμενο της παρούσας διπλωματικής εργασίας ήταν η ανάπτυξη ενός συστήματος αυτόματης γεωκωδικοποίησης ελεύθερου κειμένου, συνδυάζοντας κατάλληλα όλα τα εργαλεία εξόρυξης γεωγραφικής πληροφορίας και γεωκωδικοποίησης που είναι έχουμε στην διάθεση μας. Οι στόχοι που τέθηκαν από το αρχικό στάδιο σχεδίασης συνοψίζονται στα παρακάτω σημεία:

- Επεξεργασία πραγματικού χρόνου για το κείμενο που πληκτρολογεί ο χρήστης
- Αντιστοιχία μεταξύ των γεωγραφικών όρων του κειμένου και των δεικτών του χάρτη
- Δυνατότητα χειροκίνητης αντιστοίχισης κειμένου σε συντεταγμένες στον χάρτη
- Φιλικό και εύχρηστο γραφικό περιβάλλον εφαρμογής

Για την επίτευξη των παραπάνω στόχων, επιστρατεύτηκαν δημοφιλείς βιβλιοθήκες και εργαλεία της γλώσσας JavaScript. Παράλληλα αναπτύχθηκε ένα πλήρες σύστημα διαχείρισης κειμένου μέσω του οποίου αξιοποιούνται πλήρως όλες οι δυνατότητες της εφαρμογής Geotyping. Το σύστημα αυτό αποτελεί ένα πρότυπο παράδειγμα ενσωμάτωσης της εφαρμογής σε οποιοδήποτε σύστημα διαχείρισης περιεχομένου.

Κατά την ανάπτυξη της εφαρμογής αντιμετωπίστηκαν διάφορα ζητήματα για την λύση των οποίων επινοήθηκαν λύσεις που μπορούν να χρησιμοποιηθούν σε παρόμοιες

περιπτώσεις. Η εφαρμογή στοχεύει στην ενίσχυση της επεξεργασίας της πληροφορίας που εισάγεται καθημερινά από χρήστες στο διαδίκτυο μέσω ιστοτόπων και ιστολογίων στον τομέα της εξαγωγής γεωγραφικής πληροφορίας. Με αυτόν τον τρόπο ο κάθε χρήστης συμβάλει στην επέκταση μιας μεγάλης γεωγραφικής βάσης δεδομένων μέσα από τα κείμενα που δημοσιεύει.

8.2 Μελλοντικές επεκτάσεις

- Βελτιστοποίηση του τρόπου επεξεργασίας του κειμένου

Η τρέχουσα έκδοση της εφαρμογής λειτουργεί σε επίπεδο παραγράφου (Html block element) καθώς η αλλαγή της επιλεγμένης παραγράφου ενεργοποιεί την επεξεργασία του κειμένου της. Η επεξεργασία του κειμένου σε ένα περισσότερο λεπτομερές επίπεδο αποτελεί μια πιθανή μελλοντική επέκταση της εφαρμογής. Η εφαρμογή μπορεί να λειτουργεί σε επίπεδο προτάσεων και ιδανικά σε επίπεδο φράσεων - λέξεων όπως η υπηρεσία Google Translate. Η προτεινόμενη επέκταση προϋποθέτει την άμεση εξαγωγή των γεωγραφικών όρων του κειμένου από τον Geoparser και την ταυτόχρονη αντιστοίχιση τους σε σημεία στον χάρτη από τον Geocoder.

- Βελτιστοποίηση αποτελεσμάτων Geocoder

Λόγω της ασάφειας και των πολλαπλών νοημάτων που έχουν οι λέξεις και τα τοπωνύμια μια γλώσσας, υπάρχει πιθανότητα η διαδικασία της γεωκωδικοποίησης να επιστρέφει μη αναμενόμενα αποτελέσματα. Για την βελτιστοποίηση αυτών των αποτελεσμάτων, θα μπορούσε να χρησιμοποιηθεί η τεχνική του Viewport Biasing που προσφέρεται από την υπηρεσία γεωκωδικοποίησης της Google. Μέσω της τεχνικής αυτής, ο προγραμματιστής οριοθετεί ένα πλαίσιο στον χάρτη μέσα στο οποίο ο γεωκωδικοποιητής αναζητάει το βέλτιστο ταίριασμα φράσης και συντεταγμένων. Να σημειωθεί ότι σε περίπτωση που δεν εντοπιστεί κάτι μέσα στο πλαίσιο, η αναζήτηση συνεχίζεται και εκτός πλαισίου.

Μια δεύτερη τεχνική που θα μπορούσε να επιστρατευθεί είναι η τεχνική του Region Biasing. Ο προγραμματιστής έχει την δυνατότητα ρύθμισης της υπηρεσίας γεωκωδικοποίησης ώστε να επιστρέφει αποτελέσματα σε μια συγκεκριμένη περιοχή. Η περιοχή ορίζεται ρητά μέσω της αντίστοιχης παραμέτρου. Οι τιμές που μπορεί να πάρει η παράμετρος καθορίζονται από τον μοναδικό κωδικό της κάθε χώρας με βάση την IANA.

- Πλήρης ενσωμάτωση σε συστήματα διαχείρισης περιεχομένου

Η εφαρμογή Geotyping θα μπορούσε να ενσωματωθεί πλήρως στις δημοφιλείς πλατφόρμες διαχείρισης περιεχομένου όπως είναι το Joomla, το Drupal και το Wordpress.

9. Βιβλιογραφία

- [1] Wikimedia Foundation Inc. Website: <http://en.wikipedia.org/wiki/Geoparsing>
- [2] Wikimedia Foundation Inc. Website: <http://en.wikipedia.org/wiki/Geocoding>
- [3] Wikimedia Foundation Inc.
Website: http://en.wikipedia.org/wiki/Geographic_information_system
- [4] Wikimedia Foundation Inc.
Website: http://en.wikipedia.org/wiki/Geographic_information_retrieval
- [5] Wikimedia Foundation Inc. Website: http://en.wikipedia.org/wiki/Same_origin_policy
- [6] Wikimedia Foundation Inc. Website: <http://en.wikipedia.org/wiki/Geotagging>
- [7] Euthymios Drymonas and Dieter Pfoser. 2010. Geospatial route extraction from texts. In Proceedings of the 1st ACM SIGSPATIAL International Workshop on Data Mining for Geoinformatics (DMG '10) . ACM, New York, NY, USA, 29-37
- [8] Albert Angel, Chara Lontou, Dieter Pfoser and Alexandros Efentakis. 2008. Qualitative geocoding of persistent web pages.16th ACM SIGSPATIAL International Symposium on Advances in Geographic Information Systems. ACM-GIS 2008
- [9] Dieter Pfoser, Charikleia Lontou, Euthymios Drymonas and Savvas Georgiou. 2010. Geoblogging: user-contributed geospatial data collection and fusion.18th ACM SIGSPATIAL International Symposium on Advances in Geographic Information Systems. ACM-GIS 2010
- [10] Νικόλαος Γ. Τσιούγκος. Διπλωματική "Ανάπτυξη μιας Γεω-Microblogging εφαρμογής για Android". Ιούλιος 2010.
- [11] Γεώργιος Πρέκας. Διπλωματική "Προηγμένη Εφαρμογή Ιστού Διαχείρισης Δεδομένων Βιοεπιστημών". Δεκέμβριος 2009
- [12] Χαρίκλεια Λόντου. Διπλωματική "Ημι-αυτόματη Γεωκωδικοποίηση Ιστοσελίδων". Ιούλιος 2008
- [13] CKEditor API. Website: http://docs.cksource.com/ckeditor_api/

- [14] Google Maps API. Website: <http://code.google.com/apis/maps/index.html>
- [15] jQuery Library. Website: <http://jquery.com/>
- [16] PHP Manual. Website: <http://www.php.net/manual/en/>
- [17] MySQL Reference Manual. Website: <http://dev.mysql.com/doc/refman/5.6/en/>
- [18] Apache HTTP Server Version 2.2 Documentation. Website:
<http://httpd.apache.org/docs/2.2/en/>
- [19] JavaScript Guide. Website: <https://developer.mozilla.org/en/JavaScript/Guide>)
- [20] Browser compatibility information. Website:<http://quirksmode.org/>)
- [21] jQuery Highlight Plug-in.
Website:<http://johannburkard.de/blog/programming/javascript/highlight-javascript-text-highlighting-jquery-plugin.html>
- [22] MVC Design pattern. Website: <http://web-seminar.softlab.ntua.gr/>
- [23] Kostas Kontogiannis MVC Design pattern. <http://courses.softlab.ntua.gr/softeng/>
- [24] Usercake user management system. Website: <http://usercake.com/>
- [25] XML generator for JavaScript.
Website: <http://flesler.blogspot.com/2008/03/xmlwriter-for-javascript.html>
- [26] Yahoo Placemaker. Website: <http://developer.yahoo.com/geo/placemaker/>