



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΣΤΕΡΕΑΣ ΕΛΛΑΔΑΣ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΜΕ ΕΦΑΡΜΟΓΕΣ ΣΤΗ ΒΙΟΙΑΤΡΙΚΗ**

**Διαχείριση, επεξεργασία και οπτικοποίηση ανοικτών
διασυνδεδεμένων δεδομένων με χρήση περιβάλλοντος Wandora
και Open Street Maps**

Παπαδόπουλος Γεώργιος

**ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ
Υπεύθυνος
Αναγνωστόπουλος Ιωάννης
Επίκουρος Καθηγητής**

Λαμία, 2013

Δηλώνω υπεύθυνα ότι δεν έχω υποπέσει σε περιπτώσεις λογοκλοπής ή αντιγραφής, όπως αυτές διασαφηνίζονται παρακάτω και κατόπιν ρητών οδηγιών που έλαβα από τον επιβλέποντά μου.

Παπαδόπουλος Γεώργιος

Οδηγίες αποφυγής Λογοκλοπής και Αντιγραφής [απόσπασμα από <http://www.samos.aegean.gr/actuar/dlekkas/reports/OdigiesEPO12.pdf>]

1. Μην παραθέτετε κομμάτια βιβλίων ή άρθρων ή εργασιών άλλων αυτολεξεί χωρίς να τα περικλείετε σε εισαγωγικά και χωρίς να αναφέρετε το συγγραφέα, τη χρονολογία, τη σελίδα. Η αυτολεξεί παράθεση χωρίς εισαγωγικά χωρίς αναφορά στην πηγή, είναι λογοκλοπή. Πέραν της αυτολεξεί παράθεσης, λογοκλοπή θεωρείται και η παράφραση εδαφίων από έργα άλλων, συμπεριλαμβανομένων και έργων συμφοιτητών σας, καθώς και η παράθεση στοιχείων που άλλοι συνέλεξαν ή επεξεργάστηκαν, χωρίς αναφορά στην πηγή. Πρέπει να αναφέρετε πάντοτε με πληρότητα την πηγή κάτω από τον πίνακα ή σχέδιο, όπως στα παραθέματα.
2. Η αυτολεξεί παράθεση χωρίς εισαγωγικά, ακόμα κι αν συνοδεύεται από αναφορά στην πηγή σε κάποιο άλλο σημείο του κειμένου ή στο τέλος του, είναι αντιγραφή. Η αναφορά στην πηγή στο τέλος π.χ. μιας παραγράφου ή μιας σελίδας, δεν δικαιολογεί συρραφή εδαφίων έργου άλλου συγγραφέα, έστω και παραφρασμένων, και παρουσίασή τους ως δική σας εργασία. Αυτό τιμωρείται ως αντιγραφή.
3. Υπάρχει επίσης περιορισμός στο μέγεθος και στη συχνότητα των παραθεμάτων που μπορείτε να εντάξετε στην εργασία σας εντός εισαγωγικών. Κάθε μεγάλο παράθεμα (π.χ. σε πίνακα ή πλαίσιο, κλπ), προϋποθέτει ειδικές ρυθμίσεις, και όταν δημοσιεύεται προϋποθέτει την άδεια του συγγραφέα ή του εκδότη. Το ίδιο και οι πίνακες και τα σχέδια. Εσείς μπορείτε να χρησιμοποιείτε τέτοιο υλικό, με μέτρο, γιατί οι εργασίες είναι μικρού μεγέθους και πρέπει πάντα να κυριαρχούν οι δικές σας ιδέες.
4. Αυστηρά τιμωρείται επίσης η παρουσίαση έργου άλλων ως προσωπικής εργασίας.

Πρόλογος

Η πτυχιακή αυτή εργασία αποτελεί την ολοκλήρωση των σπουδών μου στο Πανεπιστήμιο Στερεάς Ελλάδας, τμήμα Πληροφορικής με εφαρμογές στη Βιοϊατρική και αποτελεί εργασία έρευνας και ανάπτυξης νέων εφαρμογών στα ανοιχτά δεδομένα. Περιγράφεται αναλυτικά ο ορισμός των ανοιχτών δεδομένων και τα βήματα ανάπτυξης της εφαρμογής τους σε γεωγραφικούς χάρτες.

Με την υποστήριξη του καθηγητή μου Δρ Αναγνωστόπουλου Ιωάννη αλλά και με την πολύτιμη βοήθεια του Δρ Παπαντωνίου Άγη -εδώ θα ήθελα να τους ευχαριστήσω για μία ακόμη φορά - μου δόθηκε η ευκαιρία να δω νέες και “μοντέρνες”, κατά τον δεύτερο, τεχνολογίες και να βιώσω πρωτόγνωρες εμπειρίες στον κόσμο του διαδικτύου.

Το παρόν σύγγραμμα αποτελείται από 5 κεφάλαια όπου στο πρώτο αποτελεί μία μικρή περίληψη της πτυχιακής εργασίας, στο δεύτερο γίνεται μια μικρή ιστορική αναδρομή στη γέννηση και άνθηση του διαδικτύου, καθώς και στην περιγραφή του τόσο σαν τεχνολογία όσο και σαν έννοια. Στο τρίτο κεφάλαιο περιγράφονται αναλυτικά τα εργαλεία διαχείρισης της γνώσης για τα ανοικτά διασυνδεδεμένα δεδομένα, στο τέταρτο δίνονται τα βήματα ένα προς ένα της ανάπτυξης της εφαρμογής και τέλος στο πέμπτο παρουσιάζονται συμπεράσματα και οι δυνατότητες αναβάθμισης της εφαρμογής.

Τέλος θα ήθελα να τονίσω τον χαρακτήρα της πτυχιακής εργασίας και να επισημάνω το σκοπό της. Το κυριότερο μέρος της εργασίας στηρίχτηκε στην πρακτική εφαρμογή και ανάπτυξη του κώδικα των προγραμμάτων από εμένα αλλά σε μερικά σημεία υπήρξε βιβλιογραφική και επιστημονική υποστήριξη όχι μόνο από συγγράματα και κείμενα αλλά και από forum και κοινωνικά δίκτυα. Σκοπός της εργασίας είναι να ευαισθητοποιηθούν όλο και περισσότεροι άνθρωποι και να παρέχουν εθελοντικά τα ανοιχτά δεδομένα γιατί η πρόσβαση στη γνώση πρέπει να είναι εύκολη και δωρεάν από όλους.

Η πτυχιακή αυτή εργασία
είναι αφιερωμένη
στους γονείς μου που
με στηρίζουν σε κάθε μου επιλογή.

Πίνακας περιεχομένων

Προλογος.....	2
Ευχαριστίες.....	3
Πίνακας Περιεχομένων	4
Κατάλογος Σχημάτων	6
Περίληψη.....	8
Abstract	9
ΚΕΦΑΛΑΙΟ 1 - ΕΙΣΑΓΩΓΗ.....	10
ΚΕΦΑΛΑΙΟ 2 – Το Διαδίκτυο των Δεδομένων (Web of Data).....	14
2.1 Από το Διαδίκτυο των Εγγράφων στο Διαδίκτυο των Δεδομένων	15
2.2 Ο Σημασιολογικός Ιστός και οι τεχνολογίες του.....	16
2.3 Τα Ανοικτά Δεδομένα.....	18
2.4 Οι πρωτοβουλίες της Wikipedia και της Dbpedia	20
ΚΕΦΑΛΑΙΟ 3 - Εργαλεία Διαχείρισης Γνώσης για Ανοικτά Διασυνδεδεμένα	
Δεδομένα	23
3.1 Εφαρμογές εξόρυξης, διαχείρισης και δημοσίευσης δεδομένων	24
3.2 Topic Maps	25
3.3 Η εφαρμογή Wandora	26
3.4 Τα Topic Maps στο Wandora	27
3.5 Σύντομη περιήγηση στο περιβάλλον του Wandora	30
ΚΕΦΑΛΑΙΟ 4 - Μεθοδολογία Ανάλυσης και Σχεδιασμού της εφαρμογής	37
4.1 Εισαγωγή στην SPARQL	39
4.2 Η SPARQL μέσα από το Wandora.....	40
4.3 Πρόβληματα και περιορισμοι στο Wandora	45
4.4 Η εφαρμογή OpenLayers και το OpenStreetmap	47
4.5 Εισαγωγή στα OpenStreetMaps.....	48

4.6 Εισαγωγή δεικτών στον χάρτη	49
4.7 Parsing δεδομένων στο Wandora.....	52
4.8 Κατασκευή του μενού της ιστοσελίδας και διορθώσεις	56
4.9 Δημοσιοποίηση της εφαρμογής στο διαδίκτυο μέσω του Okeanos	57
ΚΕΦΑΛΑΙΟ 5 – Συμπεράσματα και μελλοντικές επεκτάσεις.....	62
5.1 Σύνοψη της πτυχιακής εργασίας	63
5.2 Συμπεράσματα	64
5.3 Η διαφορά της εφαρμογής από τις άλλες πρωτοβουλίες	65
5.4 Μελλοντικές επεκτάσεις	67
Βιβλιογραφικές αναφορές	69
ΠΑΡΑΡΤΗΜΑ Ι - Κώδικας Προγράμματος.....	70
i. Javascript	
ii. SPARQL	
ΠΑΡΑΡΤΗΜΑ ΙΙ - Τεκμηρίωση Κώδικα Javascript	76
ΛΕΞΙΚΟ ΟΡΩΝ	78

ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ

Σχήμα	Περιγραφή	Σελίδα
2.1	Διάγραμμα σημασιολογικών συνδέσεων	15
2.2	Απεικόνιση της τριπλέτας	16
2.3	Απεικόνιση των διασυνδεδεμένων ανοιχτών δεδομένων	18
2.4	Το περιβάλλον του Virtuoso SPARQL Editor όπως φαίνεται στο διαδικτυακό τόπο http://dbpedia.org/sparql	20
3.1	Τα University και University of Central Greece είναι topics, το is dbpprop:type of είναι association και τα Thing και Public University είναι Roles	23
3.2	Απόδοση του topic map στο Wandora όπως αναφέρεται στο http://www.wandora.org/wiki/File:Topicmap_example01.gif	25
3.3	Το περιβάλλον εργασίας του Wandora	26
3.4	Το μενού επιλογών και κουμπιά συντομεύσεων για αποθήκευση εισαγωγή topic maps, RDF και OBO.	27
3.5	Συντομεύσεις για δημιουργία νέου topic, επιλογές server και άνοιγμα της R-console για R-project εφαρμογές	27
3.6	Το αριστερό παράθυρο δείχνει δύο καρτέλες, μία για τα topics και μία για αναζήτηση	28
3.7	Τα topics μπορεί να τα πάρει κατ' ευθείαν από την Dbpedia...	29
3.8	... ή από όπου αλλού θέλουμε εμείς, δηλαδή social networks, ειδήσεις RSS ή θέτοντάς του κατευθείαν ένα SPARQL ερώτημα	30
3.9	Τα ονόματα των βιομηχανιών ως topics	31
3.10	Ο πίνακας των συσχετίσεων στο κεντρικό παράθυρο	31
3.11	Topic map των κινηματογραφικών βιομηχανιών	32
3.12	Association μεταξύ των δύο βιομηχανιών	32
4.1	Μενού επιλογών	37
4.2	Παράθυρο εισαγωγής ερωτημάτων στη Dbpedia	37
4.3	Όλες οι χώρες με Hard – Rock συγκροτήματα	38
4.4	Η Ελλάδα με το συγκρότημα της και τις συντεταγμένες που έχουν δοθεί στη Wikipedia	37
4.5	Η πολιτεία του Texas με συντεταγμένες και τα συγκροτήματα	37
4.6	Ο πίνακας εμφάνισης των αποτελεσμάτων σε localhost πριν την επεξεργασία	38
4.7	Προβολή του χάρτη	15
4.8	Δείκτης που εστιάζει στο κέντρο της Αθήνας	16
4.9	Παραγωγή τυχαίων συντεταγμένων και ετικετών στο χάρτη	18
4.10	Το υλοποιημένο μενού με τέσσερα διαφορετικά ερωτήματα.	20
4.11	Εικονικός υπολογιστής είναι ανοιχτός και έτοιμος για χρήση.	23
4.12	Server settings	25
4.13	Command Window	26

5.1	Οι διαβαθμίσεις των χρωμάτων και τα ποσοστά %
5.2	Ο χάρτης μέσω του Wandora μετά την επεξεργασία
5.3	

Περίληψη

Η άνθηση και η συνεπακόλουθη ωρίμανση των τεχνολογιών του Σημαιολογικού Ιστού ή Web 3.0 όπως εναλλακτικά καλείται, βοήθησε σημαντικά στην ανάδειξη των πλέον ορθών τρόπων για τη διαχείριση, επεξεργασία και δημοσιοποίηση των δεδομένων εντός του Παγκόσμιου Ιστού. Αποτέλεσμα της τεχνολογίας αυτής είναι η ύπαρξη ενός τεράστιου σετ από ανοικτά και διασυνδεδεμένα δεδομένα. Τα δεδομένα αυτά είναι και επεξεργάσιμα από τους χρήστες του διαδικτύου, είτε από υπολογιστικές εφαρμογές, που λειτουργούν ως ευφυή λογισμικά.

Οι ερευνητικές και επιστημονικές που ασχολούνται με τις τεχνολογίες του Σημαιολογικού Ιστού καλούνται να αξιοποιήσουν τα διασυνδεδεμένα αυτά δεδομένα και να τα προσφέρουν στους χρήστες του διαδικτύου. Η επιτυχημένη αξιοποίηση, από τους χρήστες του διαδικτύου, των δεδομένων αυτών βασίζεται σε δύο πολύ σημαντικούς παράγοντες. Ο πρώτος έχει να κάνει με την κατηγοριοποίηση τους, τον έλεγχο της συνάφειας τους και την περαιτέρω επεξεργασία τους, ενώ ο δεύτερος με την οπτικοποίησή τους μέσω δυναμικών και καινοτόμων εφαρμογών.

Σκοπός της πτυχιακής αυτής εργασίας είναι αρχικά η διερεύνηση των δύο αυτών σημαντικών παραγόντων και στη συνέχεια η ανάπτυξη μιας καινοτόμας εφαρμογής οπτικοποίησης ενός σετ δεδομένων. Η επίτευξη του σκοπού της πτυχιακής αυτής εργασίας θα πραγματοποιηθεί με τη χρήση των τεχνολογιών του Σημαιολογικού Ιστού αλλά και διαμέσου της χρήσης δυο πολύ συγκεκριμένων περιβαλλόντων ανοικτού κώδικα.

Το πρώτο περιβάλλον, αφορά στην κατηγοριοποίηση, επεξεργασία και εξαγωγή διάφορων σετ δεδομένων και είναι το Wandora. Το δεύτερο που επιλέχθηκε είναι το περιβάλλον OpenStreetMap μέσω του οποίου θα πραγματοποιηθεί η οπτικοποίηση των αποτελεσμάτων του Wandora σε ένα γεωγραφικό χάρτη.

Abstract

The starting point and the subsequent maturation of Semantic Web technologies such as Web 3.0 as alternatively called, helped to highlight the best ways to manage, edit and publish data in the Web. Result of this technology is the existence of a huge set of open and interconnected data sets. These data sets are editable by users of the internet, or computer applications, operating plenty of intelligent software.

The research and scientific dealing with Semantic Web technologies, is required to build these interconnected data sets and to provide them for the users of the internet. The successful development of these data sets is based on only two, but very important factors. The first factor has to do with the classification, examination of the relevance and further processing of data sets and the second factor, with visualization through dynamic and innovative applications that will be used by users worldwide.

The purpose of this project is to first investigate these two important factors and then develop an innovative application of the data sets visualization. The purpose of this project will be achieved through the use of Semantic Web technologies but also through the use of two very specific open source environments.

The first environment use is to classify, process and export various data sets and it is called Wandora. The second environment is called OpenStreetMap which is used to place the visualization of Wandora results to a geographical map.



Κεφάλαιο 1

Εισαγωγή

Τα τελευταία χρόνια το διαδίκτυο έχει εισβάλλει σημαντικά στη ζωή μας, καθώς έχει ανοίξει νέους δρόμους τόσο στην γνώση και στην επικοινωνία, όσο και στην εκπαιδευτική διαδικασία. Μ' ένα απλό κλικ μας εισάγει σε έναν εικονικό κόσμο, μέσα στον οποίο μπορούμε να αντλήσουμε πληροφορίες για οτιδήποτε και από οποιοδήποτε σημείο του πλανήτη, να επικοινωνήσουμε με συνανθρώπους μας ακόμα κι αν βρίσκονται πολύ μακριά από εμάς, να αγοράσουμε προϊόντα, να κάνουμε συναλλαγές με τράπεζες κλπ. Έτσι γίνεται αντιληπτό πως σε λίγα χρόνια αναλφάβητος θα θεωρείται όποιος δεν μπορεί να χρησιμοποιήσει στη δουλειά του ηλεκτρονικό υπολογιστή.

Όσο περνούν τα χρόνια, τόσο περισσότεροι άνθρωποι στρέφονται στο διαδίκτυο για να αντλήσουν πληροφορίες και να αναζητήσουν κείμενα και πηγές που τους οδηγούν στην ανακάλυψη της γνώσης. Η διαδικασία άντλησης εκπαιδευτικού περιεχομένου γίνεται κατα ένα μεγάλο ποσοστό από τις μηχανές αναζήτησης όπως το Google, το Bing, τη Wikipedia και πολλές άλλες εφαρμογές αναζήτησης. Το μλονο που έχει να κάνει ο χρήστης είναι να εισάγει τη λέξη – κλειδι στο πεδίο της αναζήτησης και να πατήσει το πλήκτρο Enter. Το μεγάλο ερώτημα είναι, αν τα δεδομένα και οι κατευθύνσεις προς τη γνώση που θα επιστραφούν από τη μηχανή αναζήτησης είναι σωστά. Και να είναι τα σωστά, πάντα είναι πάρα πολλά και ίσως να μην χρειάζονται όλα ή επαναλαμβάνονται.

Το πρόβλημα αυτό μας οδηγεί στην ανάπτυξη νέων τεχνολογιών για την εξέλιξη του Παγκόσμιου Ιστού. Έτσι στις αρχές του 2007 ανακοινώνεται από την παγκόσμια ογάνωση που είναι υπεύθυνη για το διαδίκτυο ότι σιγά σιγά θα πρέπει οι σελίδες που δημιουργούνται να έχουν τέτοια δομή που θα μπορεί να γίνει κατανοητή από τον υπολογιστή και θα τον καθοδηγεί στην απόδοση της έννοιάς της. Έτσι ερευνητές και επιστήμονες συμφώνησαν, σε συνεργασία με τις κυβερνήσεις και οργανώσεις, να χτίσουν από την αρχή το διαδίκτυο, να του δώσουν έννοια και δομή και τελικά να το οδηγήσουν σε μια αναβάθμιση του Ιστού που ονομάζεται Σημασιολογικός Ιστός.

Η πτυχιακή αυτή εργασία χωρίζεται σε πέντε κεφάλαια στα οποία γίνεται κατανοητός ο όρος του διαδικτύου και του Παγκόσμιου Ιστού γενικότερα και στη συνέχεια, με τη δημιουργία μια εφαρμογής που εκμεταλλεύεται τα διασυνδεδεμένα δεδομένα κατανοούμε την τεχνολογία του και την λογική πίσω από τον κώδικα του προγράμματος.

Στο δεύτερο κεφάλαιο της εργασίας αυτής γίνεται μια ιστορική αναδρομή στο διαδίκτυο και αναλύεται πώς περναμε από το διαδίκτυο των εγγράφων και της απλής αναζήτησης πληροφορίας, στο διαδίκτυο των δεδομένων όπου η ανταλλαγή απόψεων, αρχείων και πολυμέσων καλπάζει, οι άνθρωποι όσο μακριά και να βρίσκονται επικοινωνούν, γνωρίζονται, διασκεδάζουν, ψυχαγωγούνται. Αργότερα αναλύεται το τελικό στάδιο στο οποίο έχει περιέλθει το διαδίκτυο, τον σημασιολογικό Ιστό δηλαδή. Σε αυτό το σημείο αναλύονται όροι όπως σημασιολογία και μια μια οι τεχνολογίες που χρησιμοποιούνται. Φυσικά, δεν παραλείπεται ο όρος Ανοιχτά Δεδομένα όπου δεν είναι τίποτα άλλο από την αρχική ιδέα της ανοιχτής πρόσβασης στη γνώση.

Πρωτοβουλίες όπως της Wikipedia που είναι η μεγαλύτερη on line εγκυκλοπαίδεια, να παρέχει δωρεάν και ανοιχτά όλα τα δεδομένα της, έχουν πάρει τον κόσμο με το μέρος τους και όλο και περισσότεροι βοηθούν εθελοντικά σε αυτές. Έτσι στο τέλος του κεφαλαίου αναφέρονται αναλυτικά αυτές οι πρωτοβουλίες που κάνουν την αναζήτηση και την ανακάλυψη γνώσης πολύ πιο εύκολη.

Στο τρίτο κεφάλαιο της εργασίας αυτής, γίνεται μια προσπάθεια για την επεξηγήση και κατανόηση των εργαλείων διαχείρισης της ανοιχτής γνώσης και τρόποι επεξεργασίας, και δημοσιοποίησης από και σε άλλους διαδικτυακούς τόπους τα διασυνδεδεμένα δεδομένα. Στην αρχή γίνεται αναφορά σε όλες εκείνες τις τεχνολογίες και τις εφαρμογές που χρησιμοποιούνται, η κάθε μια με τον δικό της σκοπό και έργο, και στην συνέχεια αναλύονται όροι όπως τα topic maps, πολύ σημαντικού και ισχυρού εργαλείου οπτικοποίησης της γνώσης. Αργότερα πραγματοποιείται μια πολύ γρήγορη περιήγηση σε ένα από τα 211 εργαλεία διαχείρισης των ανοιχτών δεδομένων το οποίο ονομάζεται Wandora. Σε αυτό το σημείο γίνονται κατανοητά τα βασικά κουμπιά και λειτουργίες του για την δημιουργία, επεξεργασία και δημοσιοποίηση των δεδομένων, και αναλύονται όροι όπως topic, association και occurrence, βασικά δομικά στοιχεία των ανοιχτών διασυνδεδεμένων δεδομένων. Τέλος με την υλοποίηση ενός παραδείγματος καταλήγουμε σε συμπεράσματα.

Στο τέταρτο και πιο σημαντικό κεφάλαιο της πτυχιακής εργασίας, δίνονται ένα προς ένα τα βήματα για τη δημιουργία μιας ολοκληρωμένης εφαρμογής που εκμεταλλεύεται τα ανοιχτά διασυνδεδεμένα δεδομένα που παράγονται από το διαδικτυακό τόπο της Wikipedia. Πρόκειται στην ουσία για ένα γεωγραφικό χάρτη ο οποίος αποδίδει με διαβαθμίσεις χρωμάτων ανά χώρα όποιο στατιστικό στοιχείο του έχουμε ζητήσει. Κίνητρο για τη δημιουργία αυτής της εφαρμογής αποτέλεσε ένας ελληνικός κυβερνητικός ιστότοπος, ο geodata.gov.gr ο οποίος δίνει “ανοιχτά” δεδομένα στους χρήστες με μια λίστα από διάφορα θέματα όπως γαλάζιες σημαίες ή αρχαία θέατρα και τα απεικονίζει σε ένα χάρτη. Το πρόγραμμα όμως που τρέχει από πίσω διαβάσει έτοιμα αρχεία με όλες τις πληροφορίες καταγεγραμμένες. Αναρωτηθήκαμε, γιατί να μην παίρνει τα δεδομένα κατ'ευθείαν μέσα από το διαδίκτυο αφού γίνεται λόγος για ανοιχτά δεδομένα, χωρίς την παρουσία των έτοιμων αρχείων καταγραφής; Έτσι, σε αυτό το κεφάλαιο βλέπουμε αναλυτικά την μεθοδολογία ανάλυσης και σχεδιασμού της εφαρμογής, το σχεδιασμό της και βήματα αναλυτικά, επεξηγημένα αποσπάσματα του κώδικα, τις κατηγορίες των ερωτημάτων, την γλώσσα υποβολής διαδικτυακών ερωτημάτων σε διασυνδεδεμένες βάσεις δεδομένων SPARQL και μετά την υλοποίηση και τα αποτελέσματα μαζί με εικόνες. Τέλος θα δούμε την διάθεσή της υλοποιημένης εφαρμογής στο okeanos.gnet.gr, και θα αναφερθούμε στην πρωτότυπη αυτή πρωτοβουλία αποδεικνύει ότι η χώρα μας παρά τις δύσκολες συνθήκες μπορεί να πάει τεχνολογικά πολύ μπροστά.

Στο 5ο και τελευταίο κεφάλαιο της εργασίας, αναλύονται τα συμπεράσματα και προτεινόμενες μελλοντικές επεκτάσεις της εφαρμογής αυτής καθώς επίσης και

κίνητρα για ανάπτυξη περισσότερων νέων παρόμοιων εφαρμογών. Ακολουθούν Βιβλιογραφικές αναφορές και δύο παραρτήματα με όλη τη συγγραφή του κώδικα χωρισμένα ανά κατηγορία. Η πρώτη αφορά στον κώδικα Javascript που έγραψα ενώ η δεύτερη σε κώδικα για SPARQL ερωτήματα.



Κεφάλαιο 2

Το Διαδίκτυο των Δεδομένων (Web of Data)

2.1 Από το Διαδίκτυο των Εγγράφων στο Διαδίκτυο των Δεδομένων

Το διαδίκτυο (Internet) είναι ένα δίκτυο ευρείας περιοχής (WAN) που καλύπτει όλες τις χώρες του κόσμου. Ξεκίνησε ως ερευνητικό πρόγραμμα από την κυβέρνηση των ΗΠΑ και αναπτύχθηκε αρχές της δεκαετίας του 1990. Αυτό που σήμερα ονομάζουμε Παγκόσμιος Ιστός (World Wide Web) είναι μια εφαρμογή του διαδικτύου η οποία κάνει την εμφάνισή της το 1993 και αποτελεί τη σύνθεση τριών τεχνολογιών. Αυτές είναι τα δίκτυα των ηλεκτρονικών υπολογιστών, η διαχείριση εγγράφων και πληροφοριών και τέλος το λογισμικό διεπαφής του χρήστη (User Interface). Έτσι γεννήθηκε αυτό που σήμερα ονομάζουμε πλοήγηση στο διαδίκτυο ή πιο σωστά, πλοήγηση σε διαφορετικούς τύπους δεδομένων με την υποστήριξη εγγράφων υπερκειμένου (http).

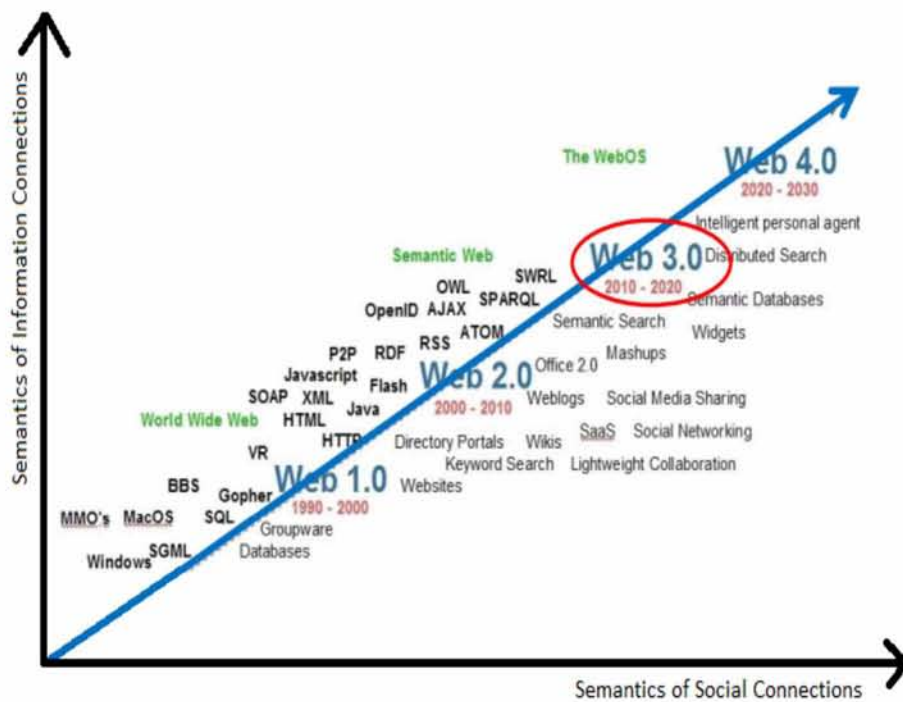
Το διαδίκτυο είναι στις μέρες μας ένα καθημερινό εργαλείο διασκέδασης, ψυχαγωγίας και πληροφόρησης. Ήταν όμως από την αρχή της δημιουργίας του έτσι όπως το βλέπουμε σήμερα; Ας κάνουμε μια ιστορική αναδρομή.

Αρχές του 1990 εμφανίστηκε ένα πρώιμο στάδιο του Παγκόσμιου Ιστού το τότε επονομαζόμενο Web 1.0. Αυτό το οποίο προσέφερε ήταν η χρήση του διαδικτύου ως διεπαφή χρήστη. Στην ουσία δεν υπήρχε η δυνατότητα να συμβάλει κάποιος στο περιεχόμενο μιας ιστοσελίδας, ούτε να γράψει κάποιο σχόλιο, αλλά μόνο να δει εικόνες και να ενημερωθεί. Η πληροφορία δεν ήταν δυναμική, ενώ μοναδικός υπεύθυνος για αλλαγή ή ενημέρωση του περιεχομένου της ιστοσελίδας ήταν μόνο ο webmaster.

Όταν παρατηρήθηκε ότι το διαδίκτυο άρχισε να γίνεται δημοφιλές και σημαντικό μέρος της καθημερινότητας, η ανάγκη για την απελευθέρωσή του ήταν επιτακτική. Έτσι, στις αρχές της χιλιετίας γεννήθηκε μια νέα μορφή Παγκόσμιου Ιστού το οποίο ονομάστηκε Web 2.0. Χαρακτηριστικό του είναι ότι οι χρήστες πλέον μπορούν να μοιράζονται μεταξύ τους πληροφορίες, έγγραφα, δεδομένα. Νέες εφαρμογές κάνουν την εμφάνισή τους όπως μηχανές αναζήτησης, blogs, ιστοσελίδες κοινωνικών δικτύων, ιστοσελίδες αναπαραγωγής πολυμέσων.

Όμως η αναζήτηση μια πληροφορίας, πόσο εύκολη μπορεί να είναι; Όταν ο χρήστης ψάχνει σε μια μηχανή αναζήτησης για κάποια πληροφορία, τα αποτελέσματα που θα του επιστρέψει είναι πάντα τα επιθυμητά;

Οδηγούμαστε λοιπόν στην ανάγκη για ανάπτυξη μιας νέας τεχνολογίας Παγκόσμιου Ιστού που θα την ονομάσουμε Web 3.0 ή Σημασιολογικός Ιστός ή Semantic Web. Προσθέτουμε δηλαδή σημασιολογικό περιεχόμενο στις ιστοσελίδες, στα αντικείμενα δηλαδή που τις περιγράφουν. Ο σημερινός αδόμητος Ιστός μετατρέπεται σε πλέγμα δεδομένων ή web-of-data. Η τεχνολογία αυτή προσφέρει δεδομένα τα οποία μπορούν να ξαναχρησιμοποιηθούν από άλλες εφαρμογές, ιστοσελίδες, άτομα, επιχειρήσεις καθώς αφού το κάθε αντικείμενο επισημαίνεται με σαφήνεια και ο υπολογιστής “κατανοεί” το τι δηλώνει. Απαραίτητη προϋπόθεση είναι τα δεδομένα αυτά να είναι ανοιχτά, όρος που θα αναλύσουμε παρακάτω.



Σχήμα 2.1: Διάγραμμα σημασιολογικών συνδέσεων.
 Πηγή: Web Semantics, Agis Papadoniou

2.2 Ο Σημασιολογικός Ιστός και οι τεχνολογίες του

Σημασιολογία είναι η διευκρίνιση των όρων που περιγράφουν κάτι. Στη γλώσσα μας κάποιοι όροι έχουν κοινή σημασία. Για παράδειγμα λέξεις όπως origin, source, principle, provenance έχουν κοινή σημασία η οποία είναι κατανοητή από τον άνθρωπο. Υπάρχουν τεχνολογίες που μπορούν να κάνουν και έναν υπολογιστή να καταλαβαίνει τους όρους και θα τις αναλύσουμε παρακάτω.

Σημασιολογικός ιστός ορίζεται ένα “δίκτυο δεδομένων το οποίο μπορούν να το επεξεργαστούν οι υπολογιστές άμεσα ή έμμεσα” σύμφωνα με τον Tim Berners-Lee, ιδρυτή του W3C. Αυτό σημαίνει δηλαδή ότι αν δοθεί σαφή έννοια για τον κάθε πόρο ή subject που πραγματεύεται μια ιστοσελίδα, τότε κατά κάποιο τρόπο ο υπολογιστής καταλαβαίνει και κατηγοριοποιεί τα δεδομένα δημιουργώντας έτσι απολύτως δομημένα στοιχεία και κατ’ επέκταση πιο γρήγορη και εύκολη περιήγηση και αναζήτηση.

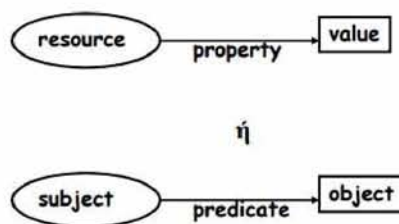
Η HyperText Markup Language (HTML) είναι η γλώσσα προγραμματισμού που χρησιμοποιείται για την κατασκευή ιστοσελίδων. Προσφέρει τη δυνατότητα απεικόνισης κειμένου, πινάκων, γραφημάτων, πολυμέσων καθώς και πληθώρα επιλογών από το χρήστη για εύκολη και ευχάριστη πλοήγηση στο διαδίκτυο.

Περιορισμός της HTML είναι το γεγονός ότι το περιεχόμενό της είναι τελείως περιγραφικό.

Έτσι κάνουν την εμφάνισή τους νέες τεχνολογίες όπως η XML ,RDF, OWL και SPARQL οι οποίες δίνουν τη σημασιολογία του κάθε αντικειμένου και δημιουργούν δομημένες ιστοσελίδες. Ας τις δούμε συνοπτικά:

-XML : Extensible Markup Language, ορίζει ένα σύνολο κανόνων για να δομηθούν σωστά τα δεδομένα της ιστοσελίδας και είναι ένα από τα πρότυπα των ανοιχτών δεδομένων. Εστιάζει στα θέματα που ορίζονται στην html και τα αναπαριστά ως δομές δεδομένων για τη χρήση τους από άλλες εφαρμογές του ιστού. Αυτό το επιτυγχάνει προσθέτοντας στα θέματα ετικέτες (tags) ως περιγραφή (description) του κάθενοσ και τελικά συνθέτει τα elements και τα attributes που αποτελούν τη λογική δομή. Τελικά προκύπτει ένα δέντρο με υπερκλάσεις και υποκλάσεις. Οι μηχανές αναζήτησης είναι πιο εύκολο και γρήγορο να ψάχνουν ετικέτες και το αποτέλεσμα επιστρέφεται με μεγαλύτερη ακρίβεια. Η σύνταξη λοιπών των δεδομένων πραγματοποιήθηκε! Η σημασιολογία πώς προκύπτει?

-RDF : Resource Description Framework, είναι ένα μοντέλο δεδομένων που ορίζει έναν κατευθυνόμενο γράφο με ετικέτες και όχι δέντρο όπως στην περίπτωση της XML. Συνήθως χρησιμοποιεί τα δομημένα δεδομένα της XML για να δώσει τη σημασιολογία τους αλλά μπορεί να δουλέψει και τελείως ανεξάρτητα. Η RDF αποτελείται από ένα σύνολο δηλώσεων. Η κάθε δήλωση είναι στην ουσία μια τριπλέτα που αποτελείται από έναν πόρο (resource) ως υποκείμενο (subject), μια ιδιότητα (property) ως κατηγορημα (predicate) και μια τιμή (value) ως αντικείμενο (object) όπως φαίνεται στο παρακάτω σχήμα. Έτσι επιτυγχάνεται η σημασιολογική τοποθέτηση του κάθε υποκειμένου.



Σχήμα 2.2: Απεικόνιση της τριπλέτας.

Πηγή: Web Semantics, Agis Papadoniou

-RDF Schema : Περιγράφει ένα σετ ομοειδών πόρων σύμφωνα πάντα με την ιδιότητα που έχουν ως προς τις τιμές , προωθώντας έτσι τα βασικά στοιχεία για την περιγραφή των οντολογιών. Αποτελεί το αρχικό στάδιο της Web Ontology Language (OWL).

-OWL : Web Ontology Language, είναι η γλώσσα σύνταξης οντολογιών στο διαδίκτυο και χρησιμοποιείται για την ανάπτυξη νέων εφαρμογών στα ανοιχτά διασυνδεδεμένα δεδομένα. Βασίζεται στην XML και στην RDFS και λίγες τροποποιήσεις έχουν γίνει σε μορφή ενημερώσεων.

-SPARQL : SPARQL Protocol and RDF Query Language, είναι η γλώσσα προγραμματισμού που εκτελεί ερωτήματα στις δομές RDF και επιστρέφει αποτελέσματα. Ακολουθεί την ίδια λογική με την γλώσσα βάσεων δεδομένων SQL, αλλά για διαδικτυακές διασυνδεδεμένες βάσεις δεδομένων. Θα αναλυθεί εκτενέστερα σε επόμενο κεφάλαιο.

2.3 Τα Ανοιχτά Δεδομένα

Όπως προαναφέρθηκε, το Web 3.0 εκμεταλλεύεται ανοιχτά δεδομένα για επαναχρησιμοποίηση από διάφορες ενδιαφερόμενες ομάδες χρηστών. Ο όρος ανοιχτά δεδομένα αναφέρεται σε σετ δεδομένων χωρίς περιορισμούς από πνευματικά δικαιώματα, πατέντες ή άλλους μηχανισμούς ελέγχου αλλά με ελεύθερη πρόσβαση από όλους. Ήδη, κυβερνητικές πρωτοβουλίες όπως data.gov έχουν αρχίσει να κερδίζουν το ενδιαφέρον όχι μόνο των πολιτών αλλά και των ερευνητών και επιστημόνων πληροφορικής για την ανάπτυξη νέων εφαρμογών.

Το υλικό πάνω στο οποίο σήμερα ισχύουν τα ανοιχτά δεδομένα, είναι χάρτες, γονίδια, ιατρικά δεδομένα και με αυτό που θα ασχοληθούμε περισσότερο είναι τα αποτελέσματα που επιστρέφει η Wikipedia που σε συνεργασία με την dbpedia.org έχει δημιουργήσει μια πολύ μεγάλη βάση δεδομένων με μεταξύ τους διασυνδεδεμένες οντολογίες ή αλλιώς mashups. Πριν όμως αναφερθούμε σε αυτήν την πρωτοβουλία της Wikipedia θα πρέπει να αναλύσουμε μερικούς πολύ σημαντικούς όρους.

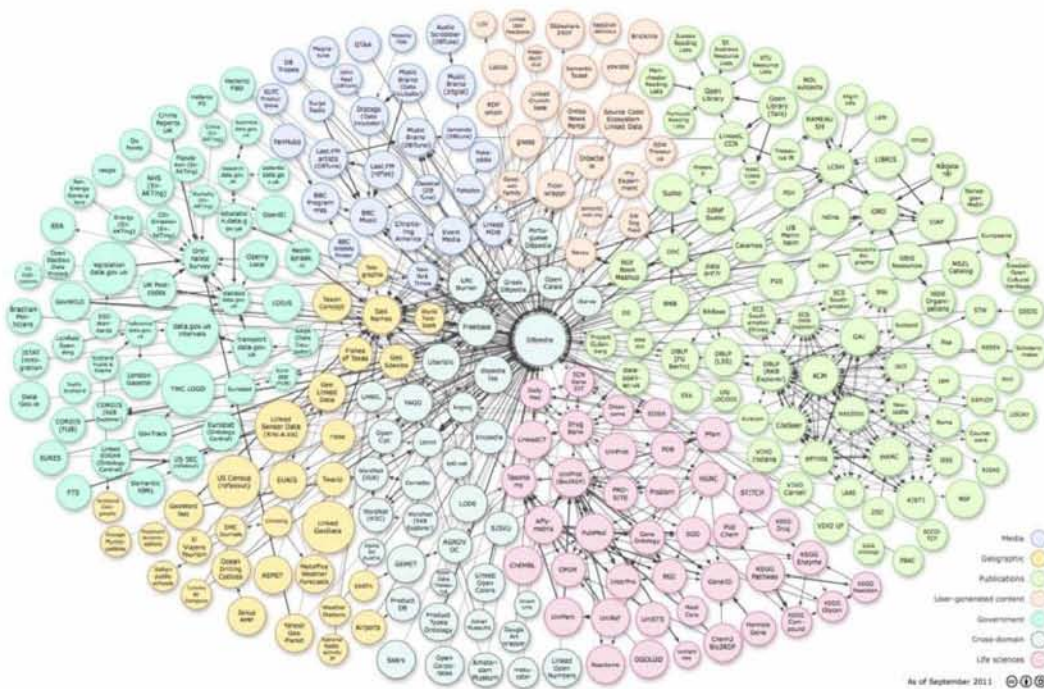
Mashup, στο web development ονομάζεται μια εφαρμογή που μπορεί και συνδυάζει και ομαδοποιεί δεδομένα με σκοπό να δημιουργήσει νέες υπηρεσίες. Συνοπτικά δηλαδή στην περίπτωση του διαδικτύου ο όρος mashup σημαίνει “μάζεμα” των ανοιχτών δεδομένων. Η τεχνική αυτή λειτουργεί ως εξής:

Όταν ο χρήστης πληκτρολογήσει το ερώτημά του σε μια μηχανή αναζήτησης, αυτό συνδέεται με πολλά άλλα δεδομένα τα οποία δεν είναι απαραίτητο να δοθούν. Τελικά

φτάνει στο σημείο να παράξει σαν αποτέλεσμα τα δεδομένα προέλευσης του ερωτήματος μέσω μιας μορφής εμπλουτισμένης αναζήτησης.

Χαρακτηριστικά, λοιπόν, του mashup είναι ο συνδυασμός, η ομαδοποίηση και απεικόνιση των δεδομένων. Τα δεδομένα τα οποία χρησιμοποιούνται είναι ήδη αρχειοθετημένα στοιχεία, έγγραφα, ιστοσελίδες που έχουν κοινά στοιχεία συνδέονται μεταξύ τους, χαρτογραφήσεις και γεωγραφικά στοιχεία και δεδομένα ροής ειδήσεων (RSS).

Η γραφική απεικόνιση και ομαδοποίηση των mashups γίνεται με τα **topic maps**. Topic map είναι ένα πρότυπο αναπαράστασης της σύνδεσης των δεδομένων. Αποτελούνται από τα topics που είναι τα αρχειοθετημένα στοιχεία και έγγραφα, τα associations που είναι οι ενώσεις και οι σχέσεις μεταξύ των topics και τα occurrences που είναι οι πληροφορίες και πηγές που συνδέονται με το κάθε topic. Έτσι δημιουργείται ένας χάρτης εννοιολογικός που εξυπηρετεί άμεσα την τεχνολογία του σημασιολογικού ιστού.



Σχήμα 2.3 : Απεικόνιση των διασυνδεδεμένων ανοιχτών δεδομένων.

Πηγή: Wikipedia

2.4 Οι πρωτοβουλίες της Wikipedia και της Dbpedia

Οι πρωτοβουλίες της Wikipedia και dbpedia αφορούν στην έναρξη της χρήσης των ανοιχτών δεδομένων σε ερευνητικό μέχρι σήμερα επίπεδο. Όπως ήδη μπορεί να γνωρίζετε, η Wikipedia περιέχει άρθρα μέσα στα οποία υπάρχουν δομημένες πληροφορίες, κατηγοριοποιήσεις και συνδέσμους προς άλλες ιστοσελίδες. Όταν ο χρήστης διαβάζει ένα άρθρο, μπορεί να μέσω μιας λέξης – κλειδί του κειμένου να βρεθεί σε μία νέα σελίδα με πληροφορίες που σχετίζονται με το άρθρο. Επίσης έχει τη δυνατότητα να προσθέσει συμπληρωματικό κείμενο ή ακόμα να γράψει ένα νέο άρθρο για κάποιο θέμα που προτείνεται από την Wikipedia, και να δημιουργήσει για αυτό νέους συνδέσμους. Με αυτόν τον τρόπο, σιγά σιγά τα δεδομένα συνδέονται μεταξύ τους και σε αυτό βοηθούν όλοι οι χρήστες του διαδικτύου. Επίσης όταν κάποιος χρήστης και συγγραφέας του άρθρου δώσει ένα σύνδεσμο από τον οποίο άντλησε τις πληροφορίες για το άρθρο αυτόματα και ο εξωτερικός σύνδεσμος διασυνδέεται με την Wikipedia, δημιουργώντας έτσι ένα πιο ευρύ φάσμα δεδομένων και πληροφοριών. Καταλαβαίνουμε λοιπόν ότι ακολουθείται μια ιεραρχία των δεδομένων και μια σειρά η οποία μπορεί ίσως να οπτικοποιηθεί σε ένα ιεραρχημένο δέντρο με κλάσεις και υποκλάσεις όπου οι υποκλάσεις δεν έχουν μόνο άλλες κάτω από αυτές αλλά συνδέονται και μεταξύ τους. Λόγω αυτής της δομής τα δεδομένα που μπορεί να ψάξει κάποιος χρήστης μπορούν να εξαχθούν πολύ πιο εύκολα και με μεγαλύτερη ακρίβεια. Αντί δηλαδή να εισάγει μια λέξη κλειδί όπως και σε μια μηχανή αναζήτησης και να επιστραφεί ένα πλήθος πιθανών αποτελεσμάτων, μπορεί να επιστραφεί ένα μόνο αποτέλεσμα και να είναι το σωστό. Και θα είναι το σωστό γιατί θα προέλθει από ένα απόλυτα δομημένο μονοπάτι.

Την πρωτοβουλία για δημιουργία ιεραρχημένων δέντρων των δεδομένων και των κλάσεων της Wikipedia την ανέλαβε η dbpedia.org. Η Dbpedia είναι μια ιστοσελίδα που ξεκίνησε το 2007 η οποία εκμεταλλεύεται την δομή της Wikipedia και καταφέρνει να εξάγει τις πληροφορίες σε μια πολύ μεγάλη βάση δεδομένων με ανοιχτή πρόσβαση σε όλους. Η βάση δεδομένων δεν είναι τίποτα άλλο από τις κλάσεις που πλέον τις ονομάζει οντολογίες, που ακολουθούν την ιεραρχημένη δομή ενός δέντρου, με τη διαφορά ότι τα κλαδιά – υποκλάσεις του διασυνδέονται και κατ' επέκταση οι οντολογίες με τη σειρά τους επίσης διασυνδέονται. Άρα οι οντολογίες της dbpedia είναι mashups και σήμερα ξεπερνούν τα 3,7 εκατομμύρια λύματα σε δεδομένα RDF ή OWL. Με την πρωτοβουλία αυτή, η Wikipedia μετατρέπεται σε μια σημασιολογική εφαρμογή που στο παρασκήνιο της υπάρχουν mashups έτοιμα να δώσουν ακριβέστατες απαντήσεις στο ερώτημα του χρήστη αλλά και τη δυνατότητα ανάπτυξης νέων ισχυρών εφαρμογών από προγραμματιστές που θα χρησιμοποιηθούν από ερευνητές και επιστήμονες όλων των κατευθύνσεων. Το όραμα είναι να ευαισθητοποιηθούν όλο και περισσότερες ιστοσελίδες, να δημιουργήσουν διασυνδέσεις έτσι ώστε η dbpedia και άλλες παρόμοιες εφαρμογές να διασυνδέσουν

όλο το διαδίκτυο. Έτσι ακόμα και οι μηχανές αναζήτησης θα μπορούν να επιστρέψουν ένα και μοναδικό σωστό αποτέλεσμα.

Αν έχουμε μια βάση δεδομένων, τότε χρησιμοποιούμε την SQL για να υποβάλουμε ερωτήματα ώστε να μας επιστραφούν τα επιθυμητά αποτελέσματα. Αν έχουμε όμως μια on – line βάση δεδομένων, τότε κατ' αντιστοιχία θα χρησιμοποιήσουμε την γλώσσα SPARQL. Εφόσον όλα τα δεδομένα εντός της dbpedia είναι γραμμένα σε βάσεις δεδομένων OWL και RDF, υποβάλλοντας SPARQL ερωτήματα, μπορούμε να εξάγουμε πίνακες αποτελεσμάτων. Τα ερωτήματα στην dbpedia υποβάλλονται σε δική της μηχανή υποβολής ερωτημάτων που ονομάζεται Virtuoso SPARQL Query Editor και φαίνεται στο σχήμα 2.4. Για την ανάπτυξη νέων εφαρμογών αλλά και για ερευνητικούς σκοπούς, τα αποτελέσματα μπορεί να τα δώσει σε διάφορες μορφές όπως HTML, CSV, Jason, XML/RDF, Javascript, ή κατευθείαν μέσα στο site της dbpedia. Για παράδειγμα μπορεί κάποιος να υποβάλει κάποια βιολογικά ερωτήματα που αφορούν γονίδια, να τα κατεβάσει σε RDF/XML και μετά να φτιάξει μια εφαρμογή που διαβάζοντας τα δεδομένα αυτά να βγάξει κάποια σημαντικά αποτελέσματα ιατρικής και στατιστικής σημασίας.

Virtuoso SPARQL Query Editor

Default Data Set Name (Graph IRI)

Query Text

```
select distinct ?Concept where {[] a ?Concept} LIMIT 100
```

(Security restrictions of this server do not allow you to retrieve remote RDF data, see [details](#).)

Results Format: (The CXML output is disabled, see [details](#))

Execution timeout: milliseconds (values less than 1000 are ignored)

Options: Strict checking of void variables

(The result can only be sent back to browser, not saved on the server, see [details](#))

Σχήμα 2.4: Το περιβάλλον του Virtuoso SPARQL Editor όπως φαίνεται στο διαδικτυακό τόπο <http://dbpedia.org/sparq>

Στο προηγούμενο σχήμα 2.4 παρατηρούμε ότι υπάρχει ένα πεδίο Default Data Set Name όπου δηλώνουμε αν θέλουμε να πάρουμε δεδομένα τα οποία υπάρχουν ήδη καταγεγραμμένα στην dbpedia.org ή μπορούμε να δηλώσουμε το live.dbpedia.org το οποίο είναι απευθείας mirror της Wikipedia για άμεσα διασυνδεδεμένα αποτελέσματα σε RDF. Ακριβώς από κάτω στο πεδίο Query Text τοποθετούμε το SPARQL ερώτημά μας και στο πεδίο Results Format επιλέγουμε σε τι format θέλουμε να γίνει η εξαγωγή των αποτελεσμάτων μας.



Κεφάλαιο 3

Εργαλεία Διαχείρισης Γνώσης για Ανοικτά Διασυνδεδεμένα
Δεδομένα

3.1 Εφαρμογές εξόρυξης, διαχείρισης και δημοσίευσης δεδομένων

Με μία γρήγορη ματιά στο διαδίκτυο μπορεί κανείς να βρει πληθώρα εφαρμογών που έχουν να κάνουν με τα διασυνδεδεμένα δεδομένα. Γενικά, υπάρχουν πάνω από 200 εφαρμογές του σημασιολογικού ιστού, η καθεμία με τον δικό της σκοπό και χρήση. Όμως λόγω του ότι μερικές εμφανίζουν κοινά στοιχεία, θα μπορούσαμε να τις χωρίσουμε σε τέσσερις διαφορετικές κατηγορίες.

Η πρώτη κατηγορία αναφέρεται σε εφαρμογές που διαβάζουν διασυνδεδεμένα αρχεία τύπου XML/RDF και βγάζουν τα αποτελέσματα απλά σε ένα πίνακα ή σε άλλες μορφές διασυνδεδεμένων αρχείων με το ίδιο όμως περιεχόμενο ανάλογα την επιθυμία του χρήστη. Σε αυτήν την κατηγορία ανήκει και το Virtuoso SPARQL Query Editor που χρησιμοποιεί η dbpedia καθώς όπως προαναφέρθηκε, η δομή της Wikipedia είναι σε μορφή XML. Στην δεύτερη κατηγορία μπορούμε να αναφερθούμε σε εφαρμογές που διαβάζουν πολύ συγκεκριμένους τύπους διασυνδεδεμένων αρχείων και συνήθως λαμβάνουν δράση σε συγκεκριμένα επιστημονικά πεδία όπως τεχνητή νοημοσύνη και συστήματα αυτομάτου ελέγχου. Αυτές συνήθως συνεργάζονται και με άλλες γλώσσες προγραμματισμού όπως Jason ή Pearl για ανάπτυξη βιολογικών και ιατρικών συνήθως εφαρμογών. Μια τρίτη κατηγορία αναφέρεται σε εφαρμογές που μπορεί κάποιος να δημιουργήσει νέες ή να επεξεργαστεί ήδη υπάρχουσες οντολογίες, mashups και αρχεία διασυνδεδεμένων δεδομένων με σκοπό τη χρήση τους από άλλα προγράμματα. Τέλος υπάρχουν εφαρμογές που είναι πολύ πιο ανεπτυγμένες, είναι γραμμένες συνήθως σε γλώσσα Java και αυτό γιατί προσφέρουν πολλές οπτικές δυνατότητες. Ο χρήστης έχει τη δυνατότητα να υποβάλλει ερωτήματα αλλά και να επεξεργάζεται στο τέλος τα αποτελέσματα, έχει την επιλογή να αποθηκεύσει, και να δημοσιοποιήσει τα δεδομένα του. Στην ουσία πρόκειται για πολύ ισχυρές εφαρμογές που συνδυάζουν και τις τρεις προηγούμενες κατηγορίες.

Σε αυτήν την πτυχιακή εργασία θα χρησιμοποιήσουμε την εφαρμογή Wandora που ανήκει στην τέταρτη κατηγορία των ισχυρών και ανεπτυγμένων εφαρμογών. Η επιλογή αυτή έγινε, καθώς το Wandora είναι ένα open source περιβάλλον το οποίο διατίθεται σε ερευνητές και προγραμματιστές ώστε να το εξελίσσουν και να δημιουργήσουν τις δικές τους εφαρμογές χωρίς να χρειάζεται να ακολουθήσουν με συγκεκριμένη λογική και βήματα που ίσως τους περιορίζουν. Επίσης παρέχει τη δυνατότητα ανάπτυξης των topic maps από τα δεδομένα και τις οντολογίες, όρος που θα αναλύσουμε πολύ συνοπτικά παρακάτω.

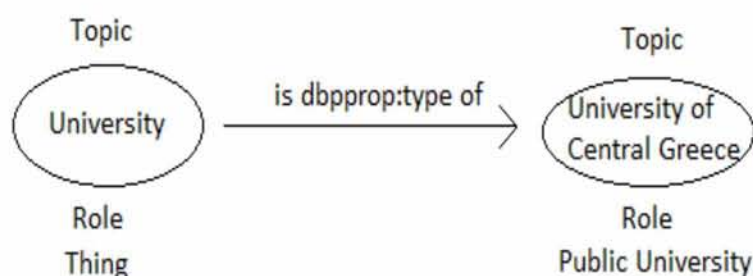
3.2 Topic Maps

Όταν ο χρήστης πραγματοποιεί μια αναζήτηση στο διαδίκτυο, δίνει λέξεις κλειδιά, τα οποία χρησιμοποιεί η μηχανή αναζήτησης ως δείκτες για να τον οδηγήσουν στην πηγή της πληροφορίας που θέλει να φτάσει. Τα topic maps αναπτύχθηκαν στις αρχές του 1990 για να εκπροσωπήσουν μια δομή τέτοια ώστε πολλαπλοί δείκτες που αναφέρονται στις διάφορες πηγές να μπορούν να συγχωνευθούν. Ωστόσο οι προγραμματιστές κατάφεραν να δημιουργήσουν μια πολύ ευρύτερη εφαρμογή προσθέτοντας στους δείκτες επιπλέον γενίκευση.

Η Wikipedia ορίζει ως topic map ένα πρότυπο για την γραφική αναπαράσταση και την ανταλλαγή δεδομένων με έμφαση στην ανιχνευσιμότητα των πληροφοριών. Ένα topic map αναπαριστά ένα θέμα (topic), τις ενώσεις (associations) και τα περιστατικά (occurences). Το θέμα μπορεί να είναι οτιδήποτε θέλουμε να περιγράψουμε. Οι σύνδεσμοι είναι οι σχέσεις μεταξύ των θεμάτων και τέλος τα περιστατικά μπορεί να είναι η πηγή των πληροφοριών που σχετίζεται με ένα συγκεκριμένο θέμα.

Το topic είναι η οντότητα στην οποία απευθύνεται κάποιος. Οτιδήποτε θελήσει ο χρήστης να περιγράψει, το θέτει ως θέμα, ως υποκείμενο. Η ένωση συνδέει δύο ή και περισσότερα τέτοια θέματα μεταξύ τους. Τα θέματα αυτά θεωρούνται οι δύο players of association. Πρέπει όμως η ένωση να οριστεί ως κάτι έτσι ώστε να γίνει κατανοητό από τον υπολογιστή του τι δηλώνει, αλλιώς δεν θα έχει νόημα η σημασιολογία. Ο ορισμός που δίνεται στην ένωση, αποτελεί με τη σειρά του και αυτό ένα topic. Επίσης πρέπει να γίνει ξεκάθαρος ο ρόλος του κάθε player. Τελικά ρόλος και player αποτελούν τα μέλη (members) της ένωσης. Το περιστατικό μπορεί να χρησιμοποιηθεί για δύο κύριους λόγους. Είτε για να συνδέσει διευθύνσεις σχετικές με το topic, είτε ως κείμενο με την περιγραφή του topic. Ορθό προγραμματιστικά θα είναι να έχουν και τα περιστατικά κάποιον ορισμό αλλά όχι ρόλους ή players.

Τα θέματα παρουσιάζουν δύο πολύ σημαντικές ιδιότητες που χωρίς αυτές το πρόγραμμα δεν θα έβγαζε τα επιθυμητά αποτελέσματα. Η πρώτη ιδιότητα είναι ότι αν δύο topic έχουν το ίδιο όνομα, μπορεί να θεωρηθούν από τον υπολογιστή ως η ίδια οντότητα, και πρέπει να πάρουν ένα εναλλακτικό όνομα. Η δεύτερη ιδιότητα είναι ότι τα topic μπορούν να έχουν μια ταυτότητα, που είναι απαραίτητη για την αποφυγή λαθών και δρα ως μοναδικό στοιχείο της οντότητας. Συνήθως η ταυτότητα αυτή είναι μια μοναδική διεύθυνση στο διαδίκτυο. Ας δούμε παρακάτω στο σχήμα 3.2 ένα απλό παράδειγμα topic map το οποίο αναφέρεται στην οντολογία Πανεπιστήμια και την σχέση της με το Πανεπιστήμιο Στερεάς Ελλάδας. Παρατηρούμε τους ρόλους που έχουν αποδοθεί στο κάθε topic, το μοναδικό όνομα και τη σχέση μεταξύ τους. Θα μπορούσε να δοθεί στο παρασκήνιο και μια ταυτότητα για το καθένα, όπου για το πρώτο θα ήταν <http://dbpedia.org/page/University> και για το δεύτερο το <http://www.ucg.gr>.



Σχήμα 3.1: Τα University και University of Central Greece είναι topics, το is dbpprop:type of είναι association και τα Thing και Public University είναι Roles.

3.3 Η εφαρμογή Wandora

Το Wandora είναι μια εφαρμογή εξαγωγής και διαχείρισης πληροφοριών από πολλές μορφές διασυνδεδεμένων βάσεων δεδομένων όπως XML/RDF ή OBO αλλά και από διαδικτυακές εφαρμογές που κάνουν χρήση των ανοιχτών διασυνδεδεμένων δεδομένων όπως Wikipedia, Dbpedia και ειδησεογραφικά κανάλια RSS. Παρέχει επίσης τη δυνατότητα στο χρήστη να επεξεργαστεί ακόμα και social networks όπως το Facebook ή το Twitter γιατί ακόμα και αυτά χρησιμοποιούν τη λογική των διασυνδεδεμένων δεδομένων έστω και εσωτερικά στους διακομιστές των διαδικτυακών τόπων τους. Το Wandora, λοιπόν, επεξεργάζεται κατάλληλα κάθε μορφής διασυνδεδεμένων δεδομένων και αποδίδει γραφικά το topic map του ερωτήματος που του θέτουμε. Φτιάχτηκε από μια μικρή ομάδα πολύ έμπειρων προγραμματιστών στη Φιλανδία και τα δύο στοιχεία που το χαρακτηρίζουν είναι η γλώσσα προγραμματισμού Java και τα topic maps.

Η εφαρμογή, παρουσιάζει ένα φιλικό και αρκετά κατανοητό στο χρήστη γραφικό περιβάλλον, παρέχει πληθώρα δυνατοτήτων όπως εισαγωγή πολλών τύπων αρχείων, απομόνωση και αποθήκευση των πληροφοριών που θέλουμε να εξάγουμε και απόδοση σε διάφορες μορφές topic maps, είτε ως πίνακας είτε ως δέντρο. Επίσης ενδιαφέρον προκαλεί το γεγονός ότι μπορεί να ανεβάσει τα topic maps σε έναν διακομιστή που φτιάχνει μόνο του, και ο απομακρυσμένος χρήστης όχι μόνο να δει τα αποτελέσματα αλλά και να κάνει την αναζήτηση νέων topics.

Επιπλέον, παρέχει τη δυνατότητα, με απλά και κατευθυνόμενα βήματα, ο ενδιαφερόμενος χρήστης, να φτιάξει νέες οντολογίες και mashups και να δομήσει έτσι τις ιστοσελίδες του και να τις διασυνδέσει με άλλες. Αν αυτό γίνει για όλες τις

ιστοσελίδες του διαδικτύου, θα δημιουργηθεί μια ευρεία ποικιλία ανοιχτών και συνδεδεμένων δεδομένων.

Επίσης, μπορεί εσωτερικά να τρέξει και άλλες πολύ ισχυρές εφαρμογές όπως η R-project, η SQL και η Java ως console αρκεί ο χρήστης να κατεβάσει τα απαραίτητα πακέτα και βιβλιοθήκες των εφαρμογών αυτών. Γενικά, η R-project είναι ένα προγραμματιστικό περιβάλλον για στατιστικούς υπολογισμούς και γραφική απεικόνιση αυτών, η SQL είναι η γλώσσα διαχείρισης και δημιουργίας βάσεων δεδομένων και η Java, είναι μια αντικειμενοστραφής γλώσσα προγραμματισμού που μπορεί να δημιουργήσει γραφικά. Το Wandora μπορεί να κάνει όλες αυτές τις εφαρμογές να δουλεύουν ταυτόχρονα και διασυνδεδεμένα, κάτι το οποίο οραματίζονται οι επιστήμονες αργότερα για το Web 4.0. Μπορεί λοιπόν κάποιος να φανταστεί τις δυνατότητες που προσφέρει το Wandora.

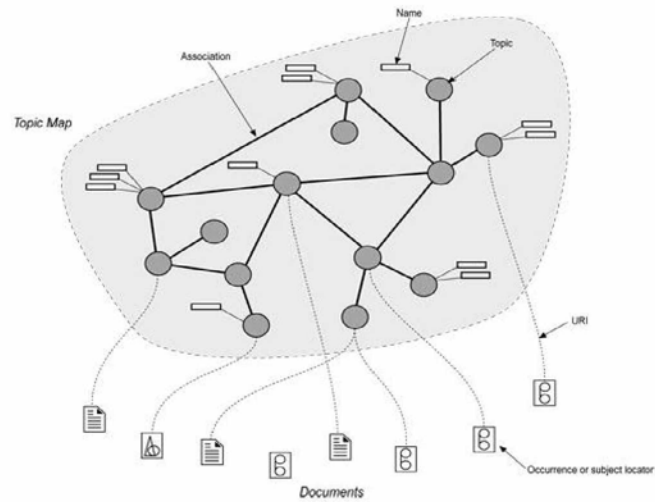
Μια πολύ ισχυρή και γνωστή εφαρμογή που έχει υλοποιηθεί είναι η απεικόνιση διάφορων στατιστικών στοιχείων μέσω ανοιχτών δεδομένων και παραλαβής αποτελεσμάτων από την R-project σε δήμους της Φιλανδίας. Η κυβέρνηση κατάφερε έτσι να ελέγχει οικονομικούς, πολιτισμικούς και κοινωνικούς παράγοντες και να αποφεύγει έτσι αρνητικές συνέπειες. Η εφαρμογή αυτή δημιουργήθηκε από την ίδια ομάδα που κατασκεύασε το Wandora, περισσότερο για να διαφημίσει μέσα από τη δουλειά αυτή τις απίστευτες δυνατότητές του.

3.4 Τα Topic Maps στο Wandora

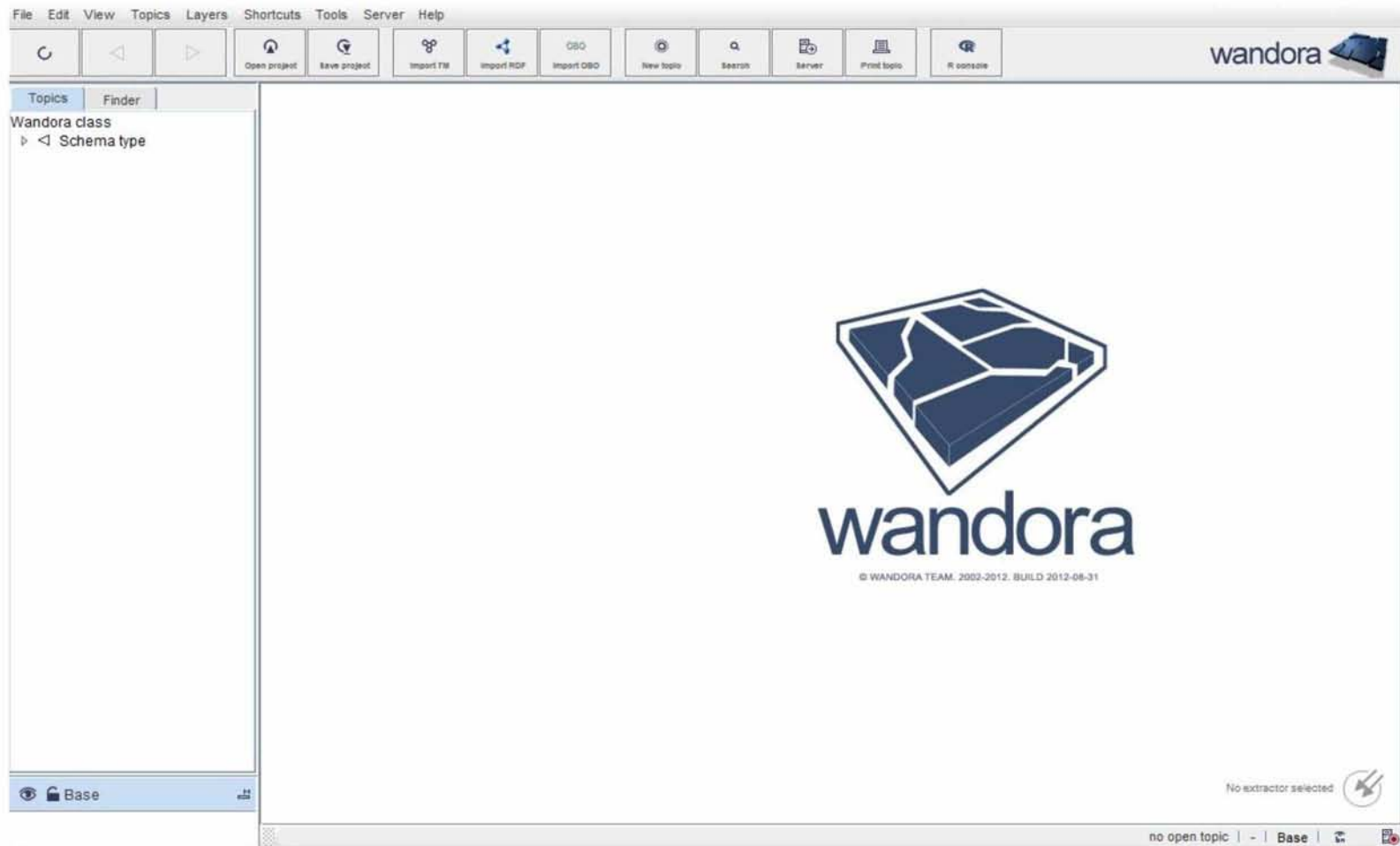
Μπορεί το βασικό στοιχείο του Wandora να είναι τα topic maps, χρησιμοποιεί όμως μια πιο ελαφριά μορφή τους και αυτό για λόγους ευκολίας και άμεσων αποτελεσμάτων, χωρίς καθυστερήσεις. Έτσι δεν είναι απαραίτητο να δοθούν πολλές λεπτομέρειες που θα κάνουν “βαρύ” το topic map. Ακολουθούν κάποιοι βασικοί κανόνες για την κατασκευή των topic maps από το wandora.

- Δεν υπάρχουν ταυτότητες για τα topics.
- Επιτρέπεται ένα μόνο όνομα για κάθε topic.
- Τα εναλλακτικά ονόματα μπορούν να μην έχουν σχέση με τα κύρια ονόματα.
- Δεν επιτρέπεται να οριστεί κάποια διεύθυνση ως περιστατικό του topic αλλά μπορεί να προστεθεί στο περιγραφικό κείμενο.
- Μόνο ένα περιστατικό επιτρέπεται για το κάθε topic.
- Μόνο ένα εναλλακτικό όνομα επιτρέπεται για το κάθε topic.
- Τα associations μπορούν να έχουν ένα μόνο player για κάθε member.
- Δεν επιτρέπονται δύο members με τον ίδιο ρόλο.

Αυτό σημαίνει ότι αν δοθεί κάποια διασυνδεδεμένη βάση δεδομένων στο wandora και περιέχει κάτι ή όλα από τα παραπάνω, αυτόματα θεωρούνται λεπτομέρειες για την ανάπτυξη διαδικτυακών εφαρμογών και διαγράφονται.



Σχήμα 3.2: Απόδοση του topic map στο Wandora όπως αναφέρεται στο http://www.wandora.org/wiki/File:Topicmap_example01.gif



Σχήμα 3.4: Το περιβάλλον εργασίας του Wandora

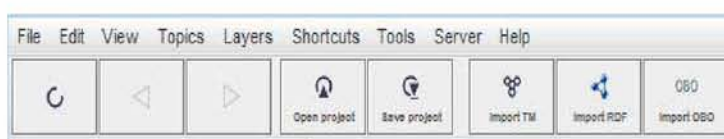
3.5 Συντομη περιήγηση στο περιβάλλον Wandora

Στο Σχήμα 3.4 απεικονίζεται η αρχική σελίδα του περιβάλλοντος του Wandora. Με μια πρώτη ματιά διακρίνει κανείς τρία βασικά στοιχεία. Το πρώτο είναι το μενού επιλογών στην κορυφή, το δεύτερο είναι το παράθυρο στα αριστερά και τέλος το κεντρικό παράθυρο που σε αυτό μετά την εξαγωγή των δεδομένων εμφανίζονται τα αποτελέσματα είτε ως πίνακας είτε ως γράφημα.

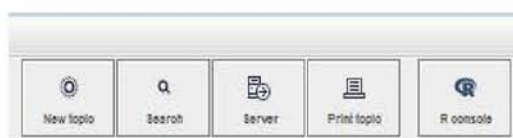
Ας κάνουμε λοιπόν μια σύντομη περιήγηση για να δούμε αναλυτικά όλες εκείνες τις λειτουργίες και δυνατότητες που μπορεί να προσφέρει στον χρήστη το περιβάλλον αυτό.

α) Η κορυφή

Στην κορυφή διακρίνουμε κουμπιά που λειτουργούν σαν συντομεύσεις των πιο βασικών λειτουργιών. Στο Σχήμα 3.5 παρατηρούμε τα κουμπιά Open Project, Save Project, τα κουμπιά εισαγωγής διάφορων τύπων αρχείων διασυνδεδεμένων δεδομένων όπως ήδη υπάρχοντα topic maps, αρχεία RDF, ακόμα και OBO. Τα τελευταία χρησιμοποιούνται συνήθως από βιολογικές εφαρμογές για καταγραφή γονιδίων και είναι σχετικά παλιό format. Στο Σχήμα 3.6 παρατηρούμε τα κουμπιά δημιουργίας νέων topic, αναζήτησης ήδη υπαρχόντων όπως επίσης και επιλογές για δημοσίευση των Project μας μέσω server, και εκτύπωση αποτελεσμάτων. Το μεγαλύτερο ενδιαφέρον παρουσιάζει το κουμπί που καλεί το User Interface της στατιστικής γλώσσας προγραμματισμού R – project. Ο χρήστης έχει τη δυνατότητα όχι μόνο να χρησιμοποιήσει τα ανοιχτά δεδομένα για εμφάνιση συσχετίσεων, αλλά να επεξεργαστεί και τα στατιστικά χαρακτηριστικά των topics, δημιουργώντας και γραφήματα και αναπτύσσοντας έτσι πανίσχυρες εφαρμογές.



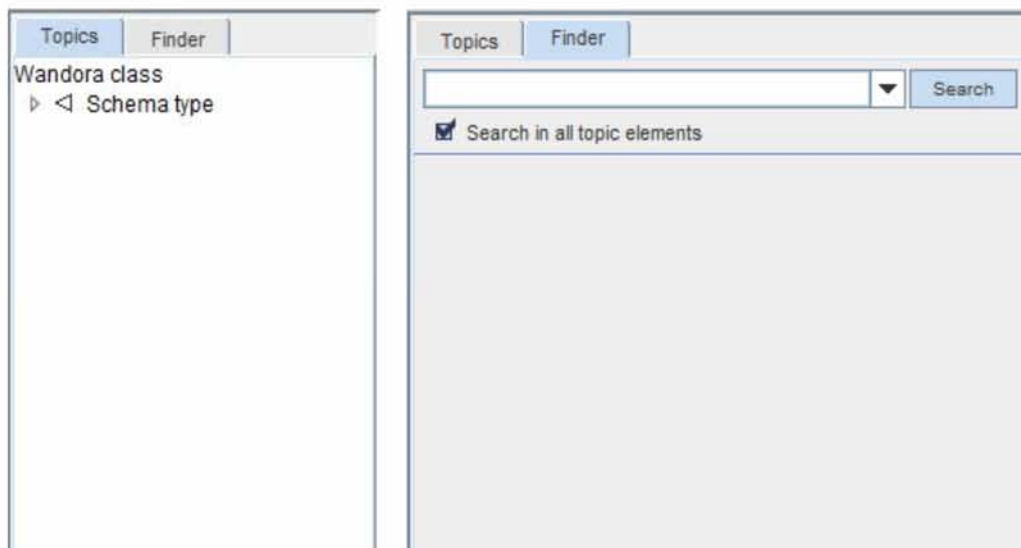
Σχήμα 3.4: Το μενού επιλογών και κουμπιά συντομεύσεων για αποθήκευση εισαγωγή topic maps, RDF και OBO.



Σχήμα 3.5: Συντομεύσεις για δημιουργία νέου topic, επιλογές server και άνοιγμα της R-console για R-project εφαρμογές.

β) Το αριστερό παράθυρο

Στο αριστερό παράθυρο της αρχικής σελίδας (Σχήμα 3.7) του Wandora, παρουσιάζονται δύο καρτέλες. Στην πρώτη καρτέλα αναφέρονται όλα όσα συνδέονται με το topic το οποίο έχουμε εισάγει, δείχνει το Schema type, τα associations και τα occurrences και ότι άλλο μπορεί να έχει να κάνει με τη διασύνδεση των δεδομένων. Ο χρήστης επιλέγοντας το βελάκι που βρίσκεται αριστερά των οντοτήτων, μπορεί να μπει και να επεξεργαστεί ότι θέλει, να αλλάξει και να εισάγει ρόλους και players, να δημιουργήσει μέσα στις οντολογίες νέα topics ή νέα στοιχεία ανάλογα με την εφαρμογή που θέλει να φτιάξει. Η δεύτερη καρτέλα αφορά στην αναζήτηση των topics που υπάρχουν και έχουμε δημιουργήσει ή επεξεργαστεί στο παρελθόν.

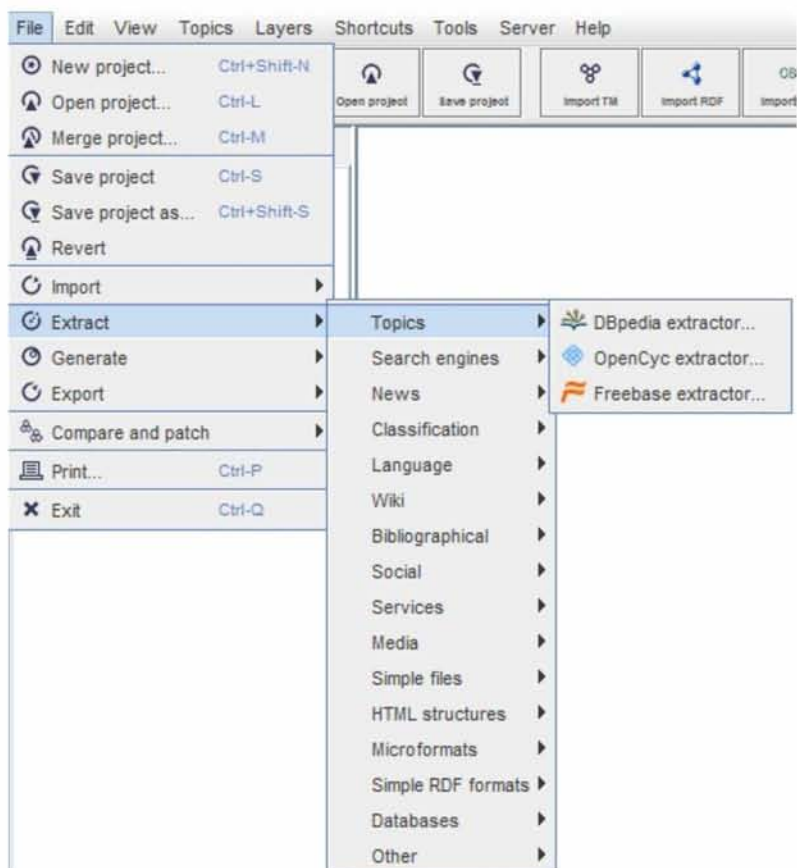


Σχήμα 3.6: Το αριστερό παράθυρο δείχνει δύο καρτέλες, μία για τα topics και μία για αναζήτηση.

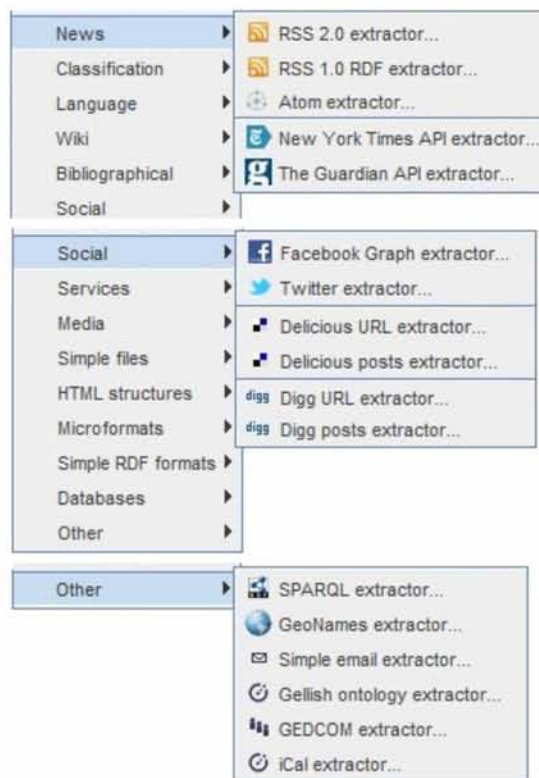
γ) Μενού επιλογών και λειτουργιών

Ας κοιτάξουμε λίγο το μενού επιλογών. Εκεί ο χρήστης του Wandora μπορεί να ακολουθήσει το μονοπάτι File → Extract και να φθάσει στο σημείο να μπορεί να επιλέξει από που θα αντλήσει τα διασυνδεδεμένα δεδομένα. Έχει λοιπόν να επιλέξει

ανάμεσα σε πολλούς διαδικτυακούς τόπους και εφαρμογές. Στο Σχήμα 3.8 παρατηρούμε τα τρία βασικότερα sites ανοιχτών δεδομένων, από τα οποία μπορεί να γίνει άμεση εξαγωγή των topics. Στο Σχήμα 3.9 παρατηρούμε πληθώρα άλλων επιλογών σε site και social networks που υποστηρίζουν τον σημασιολογικό ιστό ακόμα και ειδησεογραφικά κανάλια RSS. Στην εργασία αυτή όμως θα ασχοληθούμε αποκλειστικά και μόνο με τον SPARQL extractor και θα μιλήσουμε για αυτόν εκτενέστερα σε επόμενο κεφάλαιο.



Σχήμα 3.7: Τα topics μπορεί να τα πάρει κατ' ευθείαν από την Dbpedia..



Σχήμα 3.8: ... ή από όπου αλλού θέλουμε εμείς, δηλαδή social networks, ειδήσεις RSS ή θέτοντάς του κατευθείαν ένα SPARQL ερώτημα.

δ) Υλοποιημένο παράδειγμα

Σε αυτό το σημείο θα θέσουμε ένα πολύ απλό πρώτο ερώτημα στον SPARQL extractor του Wandora. Θα ζητήσουμε να μας επιστρέψει όλα τα ονόματα των κινηματογραφικών βιομηχανιών στον κόσμο και τις χώρες στις οποίες βρίσκονται. Παρατηρούμε ότι το παράθυρο στα αριστερά (Σχήμα 3.10) πλέον μας εμφανίζει ένα SPARQL Result Set με τα ονόματα των βιομηχανιών τα οποία είναι τα topics μας. Επίσης στο κεντρικό παράθυρο εμφάνισης των αποτελεσμάτων έχουμε τον πίνακα με τις βιομηχανίες και τις πόλεις στις οποίες βρίσκονται (Σχήμα 3.11). Για να εμφανιστεί όμως το topic map και να δούμε την γραφική απεικόνιση του ερωτήματος μας ακολουθούμε κάποια βήματα που θα δούμε παρακάτω.

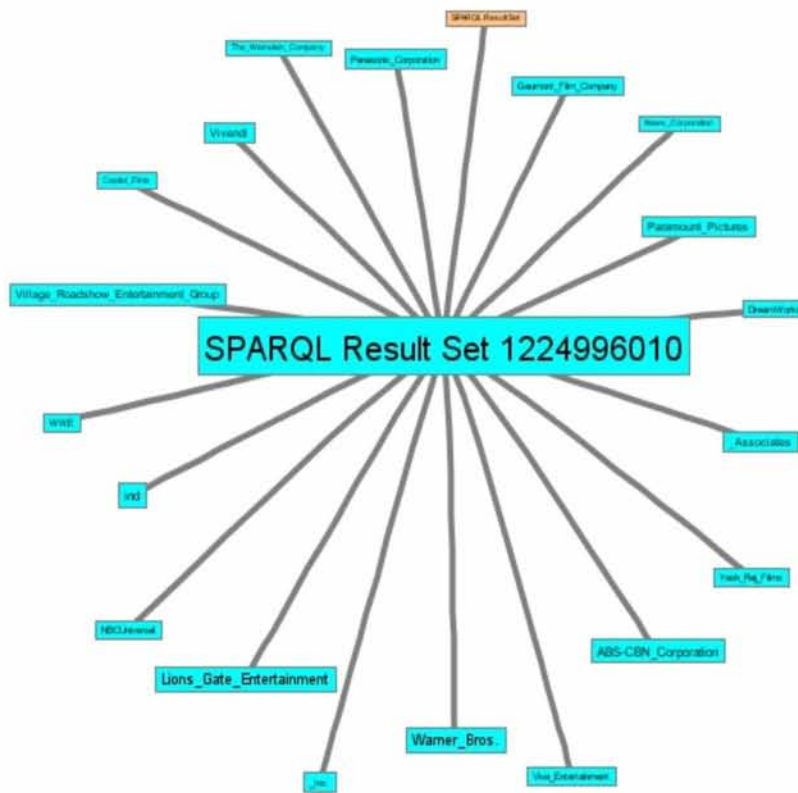


Σχήμα 3.9: Τα ονόματα των βιομηχανιών ως topics.

SPARQL Result Set 1224996010	
ind	or
Lions_Gate_Entertainment	_California
_Inc.	GMA_Network_Center
ABS-CBN_Corporation	ABS-CBN Broadcast Center, Sgt. Esguerra Ave. corner Mother Ignacia
NBCUniversal	30
NBCUniversal	GE_Building
Panasonic_Corporation	_Osaka
Vivendi	Paris
The_Weinstein_Company	New York City
Warner_Bros.	_California
_Associates	_Tokyo
Viva_Entertainment	Viva Integrated Warehouse:
WWE	Connecticut
WWE	_Connecticut
Village_Roadshow_Entertainment_Group	Melbourne, Victoria
Viva_Entertainment	Quezon City and Ortigas Center, Pasig City
Viva_Entertainment	7
News_Corporation	New York City, NY 10036
NBCUniversal	New York City 10112
Gaumont_Film_Company	Neully-sur-Seine
Paramount_Pictures	Hollywood
News_Corporation	1211
Viva_Entertainment	Viva Entertainment Head Office
_Associates	8
Yash_Raj_Films	Mumbai
Capitol_Films	Los_Angeles
Viva_Entertainment	61
_Inc.	Epifanio delos Santos corner Timog Avenues, Diliman, Quezon City
DreamWorks	Universal City, California

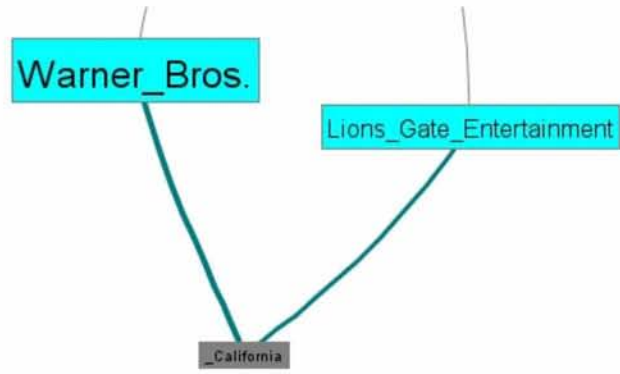
Σχήμα 3.10: Ο πίνακας των συσχετίσεων στο κεντρικό παράθυρο.

Για την εμφάνιση του topic map ακολουθούμε το μονοπάτι View → Topic Panels → Graph Topic Panel από το μενού επιλογών. Η απεικόνιση του πίνακα των αποτελεσμάτων έχει τη μορφή που φαίνεται παρακάτω στο Σχήμα 3.12.



Σχήμα 3.11: Topic map των κινηματογραφικών βιομηχανιών.

Ενδιαφέρον παρουσιάζει το παρακάτω Σχήμα 3.12 όπου βλέπουμε ότι αν επιλέξουμε το κλαδί του δέντρου με το όνομα Warner Bros τότε θα εμφανιστεί η διασύνδεση που έχει με την Lions Gate Entertainment καθώς και οι δύο βιομηχανίες βρίσκονται στην California. Η διασύνδεση δηλαδή γίνεται αυτόματα και δεν χρειάζεται κάποιος να ψάχνει με τις ώρες σε έναν πίνακα για τυχόν διασυνδέσεις.



Σχήμα 3.12: Association μεταξύ των δύο βιομηχανιών.



Κεφάλαιο 4

Μεθοδολογία Ανάλυσης και Σχεδιασμού της εφαρμογής

Αφού στο προηγούμενο κεφάλαιο είδαμε πώς λειτουργεί το περιβάλλον του Wandora εκτελώντας απλά παραδείγματα, τώρα, σε αυτό το κεφάλαιο θα συντάξουμε και θα αναλύσουμε πιο πολύπλοκα ερωτήματα που θα χρησιμοποιήσουμε στην εφαρμογή μας εξηγώντας παράλληλα το τι αυτά δηλώνουν.

Εφόσον η αρχική ιδέα της πτυχιακής αυτής ήταν να δημιουργηθεί μια εφαρμογή που να εξάγει στατιστικά δεδομένα σε έναν γεωγραφικό χάρτη, έπρεπε να βρεθεί ένας τρόπος να εκμεταλλευτούμε τα ανοιχτά δεδομένα και τις πρωτοβουλίες της Wikipedia και Dbpedia και ο χάρτης να απεικονιστεί τελικά ως Heatmap. Heatmaps ονομάζονται οι χάρτες εκείνοι που χρησιμοποιούν διαβαθμίσεις χρωμάτων, (συνήθως από ψυχρά έως θερμά χρώματα – heat) για κάποια στατιστική απεικόνιση. Τους συναντούμε συνήθως σε δελτία καιρού αλλά και σε περιπτώσεις δημοσιοποίησης άλλων μονάδων όπως δημοσιονομικών, πολιτιστικών και κοινωνικών από στατιστικές υπηρεσίες κατατάσσοντάς 'τες κατά γεωγραφική περιοχή. Εφαπτήριο αυτής της ιδέας ήταν το ελληνικό κυβερνητικό site ανοιχτών δεδομένων geodata.gov.gr το οποίο απεικονίζει διάφορα στοιχεία στον ελληνικό χάρτη όπως Αρχαία Μνημεία, Αρχαία Θέατρα, Γαλάζιες Σημείες, Νοσοκομεία, Αστυνομικά Τμήματα. Τα δεδομένα όμως για την απεικόνιση τα παίρνει από έτοιμα κατασκευασμένα αρχεία τύπου XML/RDF για γεωγραφικά δεδομένα τα οποία ονομάζονται GML. Άρα δεν κάνουμε λόγο για καθαρά ανοιχτά δεδομένα. Η ουσία είναι η λήψη live ανοιχτών δεδομένων από το διαδίκτυο οποιαδήποτε στιγμή.

Για να γίνει αυτό εφικτό χρειαζόμασταν ένα πρόγραμμα το οποίο να κάνει live εξαγωγή των ανοιχτών δεδομένων από τη βάση δεδομένων της Wikipedia, με τη χρήση της SPARQL χωρίς όμως την αποθήκευση των δεδομένων σε κάποιο εξωτερικό αρχείο XML/RDF ή OWL. Τα δεδομένα που εξάγονται, την ίδια στιγμή έπρεπε να υποστούν μια επεξεργασία για την παραγωγή των στατιστικών δεδομένων. Τέλος σημαντικό ήταν να φτιαχτεί ένα πρόγραμμα το οποίο να εμφανίζει τον χάρτη και τα αποτελέσματα μαζί.

Το πρόγραμμα το οποίο τελικά χρησιμοποιήθηκε ήταν το Wandora, γιατί πρώτον ως open source εφαρμογή, έλυne τα χέρια σε πολλά σημεία, και δεύτερον γιατί παρέχει τη δυνατότητα δημοσιοποίησης των δεδομένων σε διακομιστή. Το πρόγραμμα επεξεργασίας και παραγωγής στατιστικών δεδομένων γράφτηκε σε Javascript, γλώσσα προγραμματισμού για web development. Τέλος για την αναπαραγωγή των χαρτών χρησιμοποιήσαμε τους έτοιμους χάρτες OpenStreetMaps της εφαρμογής διαδικτυακής εφαρμογής OpenLayers.

Ο λόγος που χρησιμοποιήσαμε την εφαρμογή αυτή είναι το γεγονός ότι παρέχει δωρεάν και για ερευνητικό σκοπό χάρτες καθώς και πληθώρα παραδειγμάτων και έτοιμων εφαρμογών για γρήγορη εκμάθηση της χρήσης τους. Οι επιλογές σε χάρτες ήταν πάρα πολλές όμως χρησιμοποιήθηκαν συγκεκριμένα οι OpenStreetMaps διότι μπορούν να προβούν σε επεξεργασία εύκολα καθώς έχουν ανοιχτό κώδικα και προσβάσιμο από νέα προγράμματα και εφαρμογές. Άλλωστε χρησιμοποιούνται από πολλά Πανεπιστήμια για έρευνα και η ανάπτυξή τους είναι ταχύτερη.

Σε αυτό το κεφάλαιο, θα δούμε αναλυτικά τα βήματα της υλοποίησης όλης της πτυχιακής εργασίας, τις δυσκολίες και τη λογική που ακολούθησα πίσω από την υλοποίηση. Αρχικά θα δούμε πώς συντάσσεται ένα SPARQL ερώτημα, μετά πώς δημιουργούμε τους χάρτες σε Javascript και τέλος την ένωση των στατιστικών στοιχείων με αυτούς.

4.1 Εισαγωγή στην SPARQL

Ας αρχίσουμε με μία μικρή εισαγωγή στην SPARQL θέτοντας ένα ερώτημα στην Dbpedia. Ας πούμε ότι θέλουμε να μας επιστρέψει τα Πανεπιστήμια που υπάρχουν στον κόσμο και τις πόλεις στις οποίες βρίσκονται. Πηγαίνουμε στο διαδικτυακό τόπο της Dbpedia και τρέχουμε το ερώτημα στον Virtuoso SPARQL Extractor, εφαρμογή η οποία μας εξάγει τα δεδομένα που θέλουμε. Γράφουμε λοιπόν:

```
SELECT DISTINCT ?univ ?city
WHERE
{
?univ dbpprop:type <http://dbpedia.org/resource/University>;
dbpedia-owl:city ?city.
}
```

Επεξήγηση: Διάλεξε Πανεπιστήμιο και πόλη από το διαδικτυακό μονοπάτι που αφορά στα Πανεπιστήμια.

Το dbpprop:type και dbpedia-owl:city είναι ο τύπος δεδομένων του association που έχει δοθεί για τα topics University και City από την Dbpedia. Η επιστροφή θα είναι ένας πίνακας ο οποίος θα περιέχει όλα τα πανεπιστήμια τα οποία είναι καταγεγραμμένα στην Wikipedia. Προσοχή! Θα επιστρέψει μόνο όσα έχει η Wikipedia. Άρα, όχι όλα. Αφού στην Wikipedia μπορεί ο καθένας να γράψει ένα άρθρο, συνεπάγεται ότι κάποια στοιχεία θα είναι ελλιπή και ειδικά όταν το ερώτημά μας είναι πιο πολύπλοκο.

Ας δούμε πώς μπορεί να αξιοποιηθεί το παραπάνω ερώτημα σε ένα χάρτη. Θέτουμε το εξής ερώτημα:

```
SELECT DISTINCT ?univ ?city ?lat ?lon
WHERE
{
?univ dbpprop:type <http://dbpedia.org/resource/University>;
dbpedia-owl:city ?city;
geo:lat ?lat;
geo:long ?lon.
}
```


Επεξήγηση: Διάλεξε Πανεπιστήμιο και πόλη αλλά και τις συντεταγμένες της πόλης από το διαδικτυακό μονοπάτι που αφορά στα Πανεπιστήμια.

Θα επιστραφεί ένας νέος πίνακας με Πανεπιστήμια λιγότερα όμως από το αρχικό ερώτημα. Αυτό συμβαίνει γιατί για κάποια Πανεπιστήμια μπορεί να έχουν παραλειφθεί τα δεδομένα της πόλης ή των συντεταγμένων από τον συγγραφέα του άρθρου. Με την πάροδο του χρόνου και με την εισαγωγή όλο και περισσότερων δεδομένων στο διαδίκτυο, το πρόβλημα αυτό δεν θα υπάρχει πια.

Τελικά όμως πρέπει να φτιάξουμε ένα heatmap. Ο heatmap θα πρέπει να δείχνει διαβαθμίσεις χρωμάτων ανάλογα με το ποια χώρα έχει τα περισσότερα πανεπιστήμια (καταγεγραμμένα πάντα στην Wikipedia). Άρα χρειαζόμαστε και έναν μετρητή που θα κρατάει τον αριθμό των πανεπιστημίων ανά χώρα. Γράφουμε λοιπόν:

```
SELECT DISTINCT ?univ (COUNT(DISTINCT ?country) AS ?value)
?city ?lat ?lon
WHERE
{
?univ dbpprop:type <http://dbpedia.org/resource/University>;
dbpedia-owl:country?country;
geo:lat ?lat;
geo:long ?lon.
}
```

Επεξήγηση: Το COUNT κρατάει το πόσες φορές διαβάστηκε μια χώρα

4.2 Η SPARQL μέσα από το Wandora

Η λογική είναι απλή για κάποιον που έχει δουλέψει βάσεις δεδομένων σε SQL. Το πρόβλημα όμως είναι ότι από έναν απλό πίνακα δεν επιστρέφεται η συσχέτιση που μπορεί να έχουν τα αποτελέσματα μεταξύ τους. Εδώ λοιπόν είναι που χρησιμοποιούμε το Wandora, για να μας παράξει topic maps και να μας κάνει γραφική απεικόνιση των αποτελεσμάτων.

Ας δούμε πώς μπορεί να απεικονιστεί ένα διαφορετικό ερώτημα στο Wandora. Θέλουμε να απεικονίσουμε πόσα συγκροτήματα Hard-Rock μουσικής έχει η κάθε χώρα. Ακολουθώντας το προηγούμενο παράδειγμα σαν οδηγό γράφουμε:

```

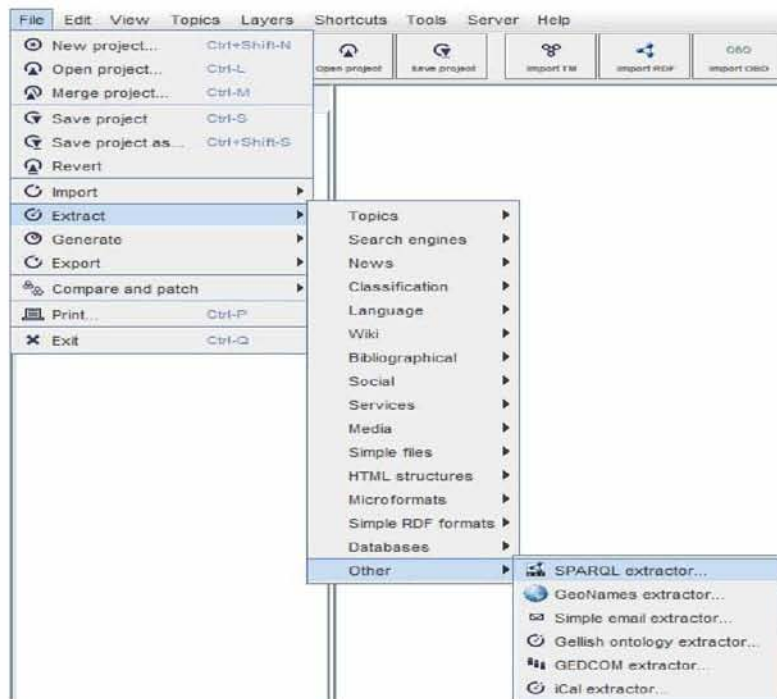
SELECT (COUNT(DISTINCT ?band) AS ?value) ?origin ?cap ?lat
?lon
WHERE
{
?aboutdbpedia-owl:genre <http://dbpedia.org/resource/Hard_rock>;
dbpedia-owl:associatedMusicalArtist ?band.
?band dbpprop:origin ?origin.
?origin dbpedia-owl:capital ?cap;
geo:lat ?lat;
geo:long ?lon.
}

```

Επεξήγηση: Βρες τον αριθμό των συγκροτημάτων της κάθε χώρας, την πρωτεύουσά της και δώσε μου τις συντετεγμένες τους.

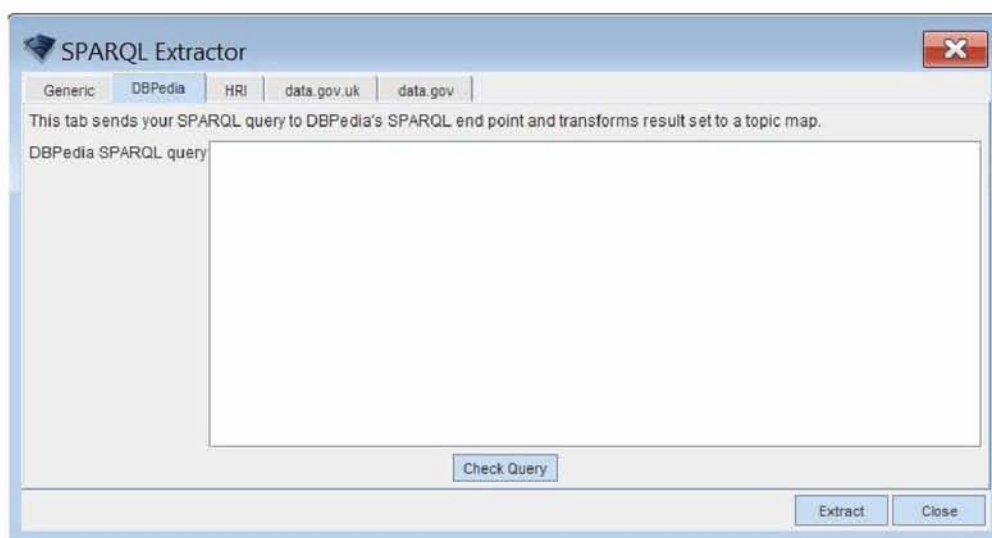
Η αναζήτηση της πρωτεύουσας για κάθε χώρα συνήθως είναι αναγκαία και αυτό γιατί τα δεδομένα όπως προαναφέραμε είναι λίγο αφηρημένα. Ο κάθε συγγραφέας μπορεί να έχει δώσει ότι συντεταγμένη θέλει αρκεί τα πέφτει μέσα στη χώρα. Οι συντεταγμένες πρωτευουσών όμως είναι πολύ συγκεκριμένες, και έτσι αποφεύγονται σημαντικά λάθη.

Πάμε να εισάγουμε το ερώτημά μας στο Wandora. Στο μενού επιλογών το μονοπάτι είναι File → Extract → Other → SPARQL Extractor.



Σχήμα 4.1: Μενού επιλογών.

Αμέσως ανοίγει το παρακάτω παράθυρο:

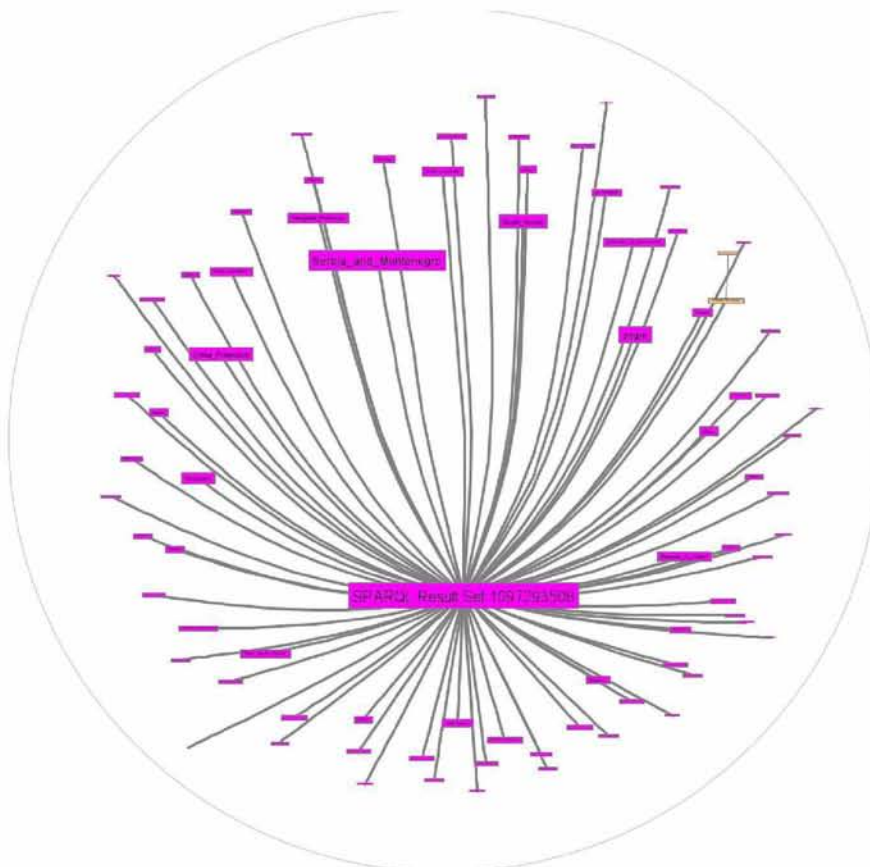


Σχήμα 4.2: Παράθυρο εισαγωγής ερωτημάτων στη Dbpedia.

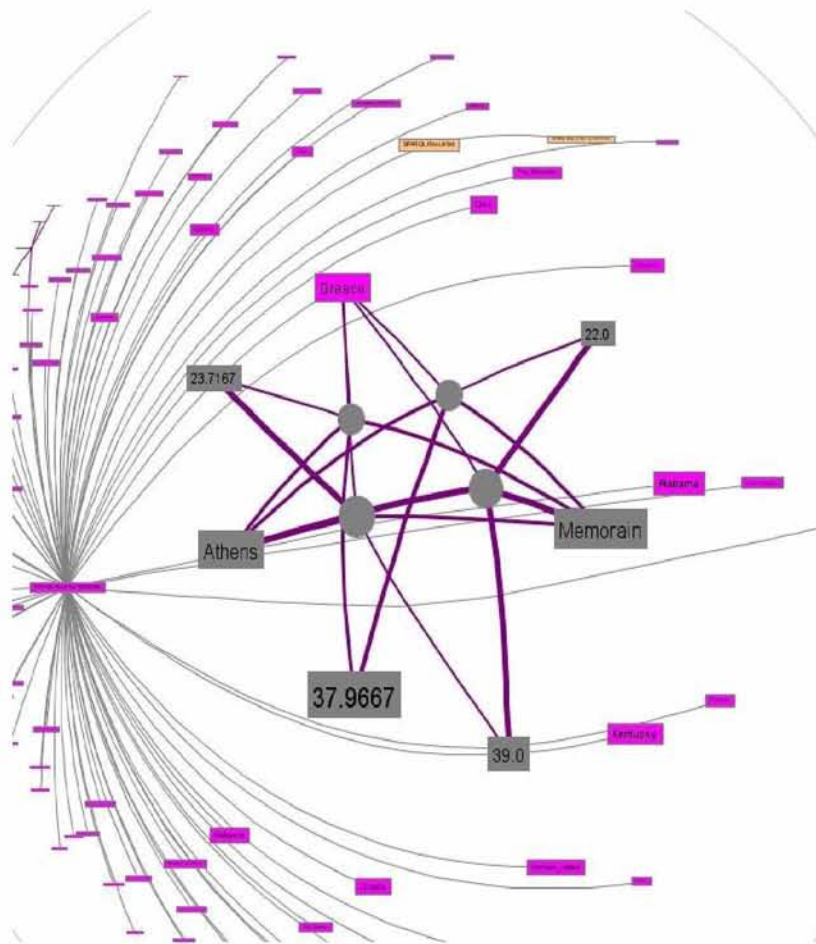
Παρατηρούμε πέντε διαφορετικές καρτέλες στον SPARQL extractor. Η πρώτη αφορά σε ερωτήματα που μπορούν να γίνονται σε ένα τοπικό δίκτυο ανοιχτών δεδομένων, η δεύτερη σε ερωτήματα στην Dbpedia και οι τρεις τελευταίες σε κυβερνητικούς ιστότοπους που υποστηρίζουν ανοιχτά δεδομένα. Για το παράδειγμά μας θα χρησιμοποιήσουμε την δεύτερη καρτέλα για ερωτήματα στην Dbpedia. Εισάγουμε το ερώτημα και επιλέγουμε extract. Προαιρετικά, αν δεν έχουμε δει ότι δουλεύει σωστά στο Virtuoso, ελέγχουμε για την ορθότητα του κώδικά μας πατώντας το Check Query.

Τα αποτελέσματα βγαίνουν και είναι ένας πίνακας όπως ακριβώς είδαμε και σε προηγούμενο κεφάλαιο. Θέλουμε όμως να δούμε τις συσχετίσεις και το topic map. Ακολουθούμε το μονοπάτι όπως έχουμε δει το οποίο είναι το View → Topic panels → Graph topic panel. Το αποτέλεσμα είναι εντυπωσιακό καθώς το ερώτημα που θέσαμε έχει πάρα πολλές εγγραφές και το topic map είναι πολύ μεγάλο με πολύ περισσότερα δεδομένα στο κάθε φύλλο του δέντρου και φυσικά δεν έχει καμία σχέση με την απλή προβολή σε πίνακα και φαίνεται στα παρακάτω Σχήμα 4.3.

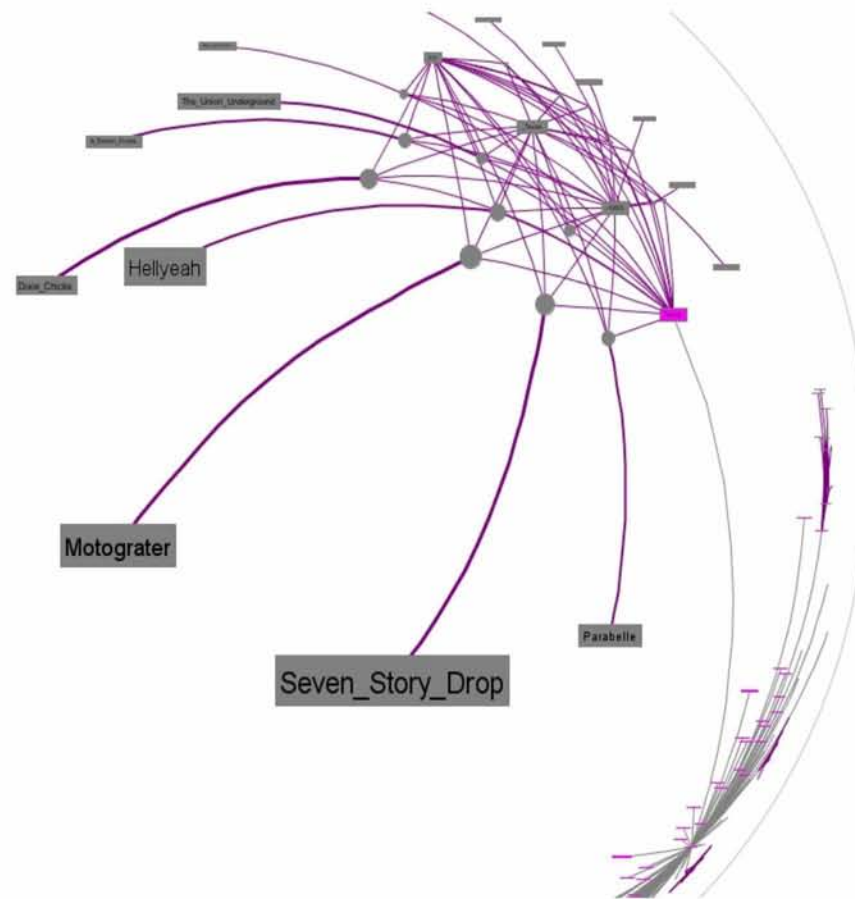
Στα παρακάτω Σχήματα 4.3, 4.4 και 4.5 φαίνεται καθαρά η διασύνδεση των δεδομένων. Στο Σχήμα 4.3 απεικονίζονται όλα τα αποτελέσματα συνδεδεμένα σε ένα κεντρικό τοπίο το οποίο αποτελεί το σύνολο των αποτελεσμάτων που στην περίπτωσή μας είναι όλες οι χώρες με τα συγκροτήματά τους. Αν επιλέξουμε ένα από αυτά τα αποτελέσματα, τότε θα απεικονιστούν όλες οι διασυνδέσεις που έχουν να κάνουν με το ερώτημα το οποίο θέσαμε. Αν παρατηρήσουμε στο Σχήμα 4.4 η Αθήνα, το συγκρότημα και οι συντεταγμένες σχηματίζουν ένα πλέγμα, δείχνοντάς μας ότι αυτά είναι μοναδικά, και διασυνδέονται μεταξύ τους μόνο. Στο Σχήμα 4.5 απεικονίζεται ένα ακόμα πιο πολύπλοκο παράδειγμα, αυτό της πολιτείας του Texas, η οποία έχει παραπάνω από ένα συγκροτήματα οπότε τα associations είναι πολύ περισσότερα σχηματίζοντας ένα μεγαλύτερο πλέγμα. Πρέπει να τονίσουμε ότι το γραφικό αυτό περιβάλλον μπορεί να υποστεί επεξεργασία σε γλώσσα προγραμματισμού Java και να απεικονίσει έναν χάρτη ή ότι άλλο θέλουμε δημιουργώντας έτσι πολύ ισχυρές stand alone εφαρμογές. Το μοναδικό πρόβλημα είναι ότι το γραφικό αυτό περιβάλλον που θα δημιουργηθεί από το χρήστη περιορίζεται μόνο στον υπολογιστή του και δεν μπορεί να αποδοθεί στο διαδίκτυο.



Σχήμα 4.3 : Όλες οι χώρες με Hard – Rock συγκροτήματα



Σχήμα 4.4: Η Ελλάδα με το συγκρότημα της και τις συντεταγμένες που έχουν δοθεί στη Wikipedia



Σχήμα 4.5: Η πολιτεία του Texas με τις συντεταγμένες και τα συγκροτήματα της

4.3 Πρόβλημα και περιορισμοί στο Wandora

Μπορεί το Wandora να προσφέρει πολλές δυνατότητες τόσο στην επεξεργασία όσο και στην απεικόνιση των ανοιχτών και διασυνδεδεμένων δεδομένων, υστερεί όμως λίγο στην δημοσιοποίηση τους μέσω του διακομιστή του. Να σημειωθεί βέβαια ότι γίνεται πολύ μεγάλη προσπάθεια για την εξάλειψη αυτών των προβλημάτων αφού είναι open source πρόγραμμα και κερδίζει όλο και περισσότερους προγραμματιστές και ερευνητές για την ανάπτυξη των δικών τους εφαρμογών. Ας δούμε αυτούς τους περιορισμούς.

Όπως προαναφέρθηκε, το Wandora δεν μπορεί να αποδώσει γραφικά τριών διαστάσεων στο διαδίκτυο. Τα αποτελέσματά μας δίνονται σε μορφή πίνακα όπως ακριβώς δηλαδή και στο Virtuoso SPARQL της Dbpedia ή σε δέντρο σε δύο διαστάσεων χωρίς όμως την δυναμικότητα και τις δυνατότητες που δίνονται στην stand – alone εφαρμογή. Αυτό συμβαίνει γιατί η Java δεν μπορεί να διαβαστεί από ένα διαδικτυακό τόπο γραμμένο σε απλή html. Περιορισμός στο θέμα γραφικών όμως είναι η δυναμικότητα. Ο χρήστης δεν έχει τη δυνατότητα να επιλέξει ένα φύλλο και να του αποδώσει όλες τις διασυνδέσεις. Τα φύλλα όλα είναι ανοιχτά και ανεπτυγμένα στο χώρο σε δύο διαστάσεις, που δεν αποτελεί και την πιο εύκολη λύση στον απομακρυσμένο χρήστη. Είναι όμως πολύ καλύτερα από την απόδοση μόνο του πίνακα με τα αποτελέσματα. Στην περίπτωση της παρούσας πτυχιακής εργασίας πρέπει να αλλάχθει σε πάρα πολλά σημεία ο κώδικας που τρέχει ο διακομιστής του Wandora. Η HTML που αποδίδει τον πίνακα θα υποστεί επεξεργασία και εσωτερικά θα γραφτεί ένα κομμάτι Javascript το οποίο θα μπορεί να αποδώσει μια τα αποτελέσματα των ερωτημάτων σε γεωγραφικούς χάρτες.

Ένας ακόμα περιορισμός και πιστεύω σημαντικό πρόβλημα του Wandora είναι ότι δεν μπορεί ένας χρήστης από απόσταση να θέσει στον διακομιστή ερώτημα SPARQL και να του επιστρέψει αποτέλεσμα live. Η υποβολή SPARQL ερωτημάτων είναι μέχρι στιγμής καθαρά θέμα του developer που διαχειρίζεται το Wandora μέσα από κάποιον δημόσιο διακομιστή. Στην παρούσα πτυχιακή εργασία έχουν υποβληθεί ήδη κάποια ερωτήματα τα οποία απεικονίζονται και είναι προσβάσιμα από απομακρυσμένους χρήστες χωρίς τη δυνατότητα επεξεργασίας τους ή αναζήτησης περισσότερων πληροφοριών. Το πρόβλημα αυτό μπορεί να λυθεί αφού αποτελεί αντικείμενο μελλοντικής έρευνας.

Στο επόμενο Σχήμα 4.6 βλέπουμε την εικόνα της απόδοσης των αποτελεσμάτων σε πίνακα, όπως φαίνεται στον τοπικό διακομιστή της εφαρμογής Wandora.

SPARQL Result Set 128087567
<http://www.wandora.org/sparql/resultset/128087567>

SPARQL Result Set

Associations of type ⁽¹²⁴⁾

SPARQL Result Set 128087567				
cap	lat	lon	origin	value
Yokohama	35.4475	139.642	Kanagawa_Prefecture	1
Sapporo	43.0	142.0	Hokkaido	1
Lima	-12.0433	-77.0283	Peru	1
Bucharest	44.4167	26.1	Romania	2
_Alabama	33.0	-86.6667	Alabama	3
Skopje	41.6	21.4333	Republic_of_Macedonia	1
Canberra	-35.3	149.133	Australia	37
Sarajevo	43.8667	18.4167	Bosnia_and_Herzegovina	4
Copenhagen	56.0	12.5667	Denmark	9
_New_York	43.0	-75.0	New_York	57
Warsaw	52.2167	21.0333	Poland	1
Skopje	42.0	21.7	Republic_of_Macedonia	1
Dublin	53.3442	-6.2675	Republic_of_Ireland	5
Riga	57.0	25.0	Latvia	1

4.4 Η εφαρμογή OpenLayers και το OpenStreetmap

Τα Open Layers είναι μια βιβλιοθήκη της JavaScript που χρησιμοποιείται για την εμφάνιση δεδομένων σε χάρτη, σε πιο σύγχρονα προγράμματα περιήγησης στον ιστό, χωρίς την εξάρτηση από κάποιον διακομιστή. Αυτό σημαίνει ότι μπορούν να σταθεί και μόνη της χωρίς την εξάρτηση της σύνδεσης στο διαδίκτυο. Τα Open Layers μέσω της JavaScript API δημιουργούν γεωγραφικές εφαρμογές, παρόμοιες με τα Google Maps με μια όμως πολύ σημαντική διαφορά. Τα OpenLayers είναι ένα ελεύθερο λογισμικό, που αναπτύχθηκε από προγραμματιστές των ανοιχτών δεδομένων.

Μία από τις βασικές αρχές των Open Layers για την ανάπτυξη του κώδικα είναι ότι τα παραδείγματα που κάθε φορά υλοποιούνται, όσο μικρά και να είναι, διατηρούνται σε ένα ανοιχτό forum προσβάσιμο από όλους έτσι ώστε κάθε φορά που δημιουργείται κάτι, να υπάρχει αποθηκευμένο και έτοιμο για χρήση, χωρίς να χρειαστεί να φτιαχτεί από την αρχή. Έτσι διατηρούνται σειρές από παραδείγματα με τις περισσότερες λειτουργίες εμπλουτίζοντας έτσι τη βιβλιοθήκη και κατ' επέκταση τη δύναμη της εφαρμογής. Τα παραδείγματα που παρέχονται μέχρι στιγμής είναι περισσότερα από 100 και είναι ανοιχτά σε όλους για την δημιουργία δυναμικών εφαρμογών.

Το OpenStreetMap (OSM) είναι ένας από τους χάρτες που παρέχουν τα OpenLayers, είναι χάρτης ανοιχτού κώδικα και επεξεργάσιμος από όλο τον κόσμο. Η ιδέα ξεκίνησε από τον Steve Coast ο οποίος χρησιμοποιούσε τα δεδομένα που καταγράφονταν από τα συστήματα GPS καθώς και πολλά ανοιχτά γεωγραφικά δεδομένα όπως αεροφωτογραφίες, συντεταγμένες και σύνορα πόλεων καθώς και πολλά στατιστικά δεδομένα. Η ιδέα άρεσε στον κόσμο και μέχρι σήμερα πάνω από 1 εκατομμύριο άνθρωποι στον κόσμο συμμετέχουν εθελοντικά σε αυτήν την πρωτοβουλία. Η ιστοσελίδα του υποστηρίζεται από τον OpenStreetMap Foundation, ένα μη κερδοσκοπικό οργανισμό που εδρεύει στην Αγγλία.

Τα στοιχεία τα οποία συνεχώς προσθέτονται, αναβαθμίζονται σε όλους τους χρήστες των χαρτών αυτών και είναι διαθέσιμα σε όλους όπως ακριβώς και τα ανοιχτά δεδομένα. Σκοπός είναι η ενθάρρυνση για ανάπτυξη και διανομή των δωρεάν γεωχωρικών δεδομένων. Επίσης παρέχει στοιχεία επισκόπησης της γης για οποιονδήποτε να χρησιμοποιήσει και να τα μοιραστεί. Ενδιαφέρον παρουσιάζει το γεγονός ότι όλο και περισσότεροι άνθρωποι θέλησαν να βοηθήσουν. Χαρακτηριστικό είναι ότι στις αρχές της εφαρμογής, μια ομάδα ποδηλατών χαρτογράφουσε με το GPS τα ποδηλατικά μονοπάτια, και άλλοι με ιστιοπλοϊκά, τα θαλάσσια μονοπάτια, της Μεγάλης Βρετανίας. Όλα σχεδόν τα δεδομένα εισήχθησαν από μετρητές, φωτογραφίες και αεροφωτογραφίες που ανήκαν σε εθελοντές.

Σήμερα πολλά Πανεπιστήμια στηρίζουν τις έρευνες και τα προγράμματα τους στο OpenStreetMap, πολλοί είναι οι φοιτητές που αναπτύσσουν πάνω σε αυτό το χάρτη συνεχώς νέες εφαρμογές και εισάγουν συνεχώς νέα δεδομένα και στατιστικά

και γεωλογικά στοιχεία. Ακόμα και οι κυβερνήσεις για την παραγωγή και δημοσιοποίηση ανοιχτών δεδομένων χρησιμοποιούν το OpenStreetMap ως βασική εφαρμογή γεωγραφικών στοιχείων.

4.5 Εισαγωγή στα OpenStreetMaps

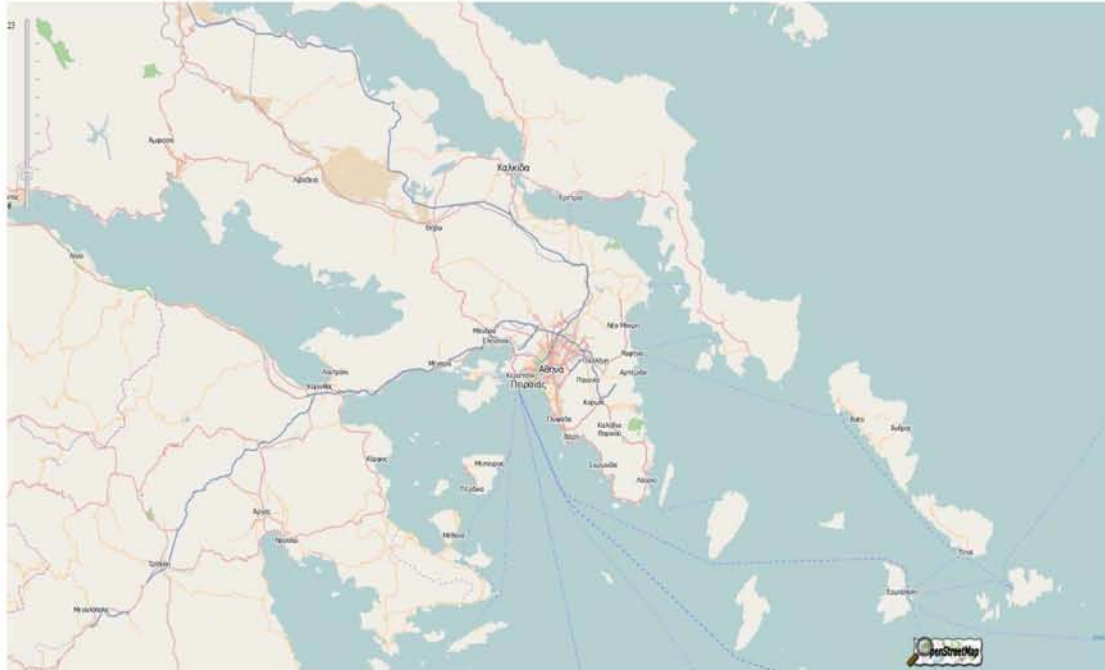
Σε αυτό το σημείο θα δούμε πώς μπορούμε να εμφανίσουμε ένα χάρτη σε μια προσωπική ιστοσελίδα. Πρόκειται για ένα απλό javascript το οποίο το εισάγουμε στην ιστοσελίδα και φαίνεται παρακάτω

```
<script type="text/javascript">
  var map, layer;
  function init(){
    map = new OpenLayers.Map( 'map');
    layer = new OpenLayers.Layer.OSM( "Simple OSM Map");
    map.addLayer(layer);
    map.setCenter(
      new OpenLayers.LonLat(465675.57551, 4201372.41529 ).transform(
        new OpenLayers.Projection("EPSG:4326"),
        map.getProjectionObject()
      ), 12
    );
  }
</script>
```

Ενδιαφέρον παρουσιάζει και είναι πολύ σημαντικό το Projection των συντεταγμένων. Το πρόβλημα είναι ότι κάθε χώρα έχει δικό της σύστημα απεικόνισης συντεταγμένων (ESPG) και έχουν οριστεί δύο global. Το ένα είναι το WGS84- ESPG: 4326 το δεύτερο το Spherical mercator ESPG:900913 το οποίο το έχει υιοθετήσει η Google για τους χάρτες της. Στην Ελλάδα χρησιμοποιούμε το GGRS – ESPG:2100. Για την αλλαγή των συντεταγμένων είναι αναγκαίο ένα ακόμα javascript για την μετατροπή, το οποίο δίνεται από τα γεωλογικά ινστιτούτα της κάθε χώρας. Όλοι οι μετατροπείς συντεταγμένων είναι αποθηκευμένοι σε μια οργάνωση στις ΗΠΑ, την Proj4js.org. Αν δηλαδή το κέντρο του χάρτη μας θέλουμε να το ορίσουμε σε συντεταγμένες που δίνονται με το ελληνικό σύστημα πρέπει σύμφωνα με την οργάνωση αυτή να εισάγουμε το παρακάτω javascript.

```
<script type="text/javascript">
Proj4js.defs["EPSG:2100"] = "+proj=tmerc +lat_0=0 +lon_0=24 +k=0.9996
+x_0=500000 +y_0=0 +ellps=GRS80 +towgs84=-199.87,74.79,246.62,0,0,0,0
+units=m +no_defs";
</script>
```

Μετά το transform του κώδικα μπορεί να αναγνωρίσει τις συντεταγμένες και να τις μετατρέψει αυτόματα σε global, στο σύστημα δηλαδή WGS84. Στο παράδειγμα έχουν δοθεί οι συντεταγμένες της Αθήνας στο ελληνικό σύστημα συντεταγμένων. Το αποτέλεσμα του κώδικα φαίνεται στο παρακάτω Σχήμα 4.6.



Σχήμα 4.7: Προβολή του χάρτη

4.6 Εισαγωγή δεικτών στον χάρτη

α) Εισαγωγή ενός μόνο δείκτη

Έστω τώρα ότι δεν θέλουμε να ορίσουμε μόνο κέντρο στο χάρτη αλλά να δώσουμε και μια συντεταγμένη που θα κάνει εμφανές ένα σημείο ενδιαφέροντος. Θα πρέπει λοιπόν να εισάγουμε με κάποιο τρόπο μια πινέζα, έναν δείκτη δηλαδή του σημείου που θέλουμε να απεικονίσουμε. Αυτό γίνεται προσθέτοντας στον παραπάνω κώδικα το εξής:

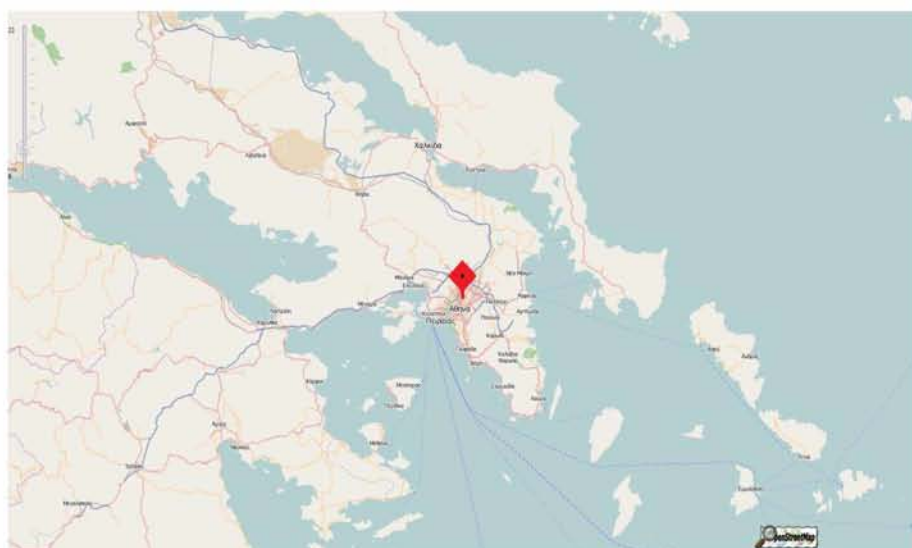
```
var markers = new OpenLayers.Layer.Markers( "Markers" );  
map.addLayer(markers);  
var size = new OpenLayers.Size(21,25);  
var offset = new OpenLayers.Pixel(-(size.w/2), -size.h);
```

```

var icon = new
OpenLayers.Icon('http://www.openlayers.org/dev/img/marker.png',size,offset);
markers.addMarker(new OpenLayers.Marker(new
OpenLayers.LonLat(465675.57551, 4201372.41529 ).transform(
new OpenLayers.Projection("EPSG:4326")),icon));

```

Γνωστοποιούμε δηλαδή στο πρόγραμμα ότι θα γίνει εισαγωγή ενός δείκτη, ορίζουμε το μέγεθος που θα έχει, δίνουμε το μονοπάτι που θα βρει την εικόνα της πινέζας και τέλος δίνουμε τις συντεταγμένες στις οποίες θα το εισάγει. Παρατηρείστε ότι έχουμε βάλει ως σημείο ενδιαφέροντος το κέντρο της Αθήνας, χρησιμοποιώντας ξανά τον μετατροπέα συντεταγμένων. Το συγκεκριμένο παράδειγμα θα έχει σαν αποτέλεσμα το επόμενο Σχήμα 4.7:



Σχήμα 4.8: Δείκτης που εστιάζει στο κέντρο της Αθήνας.

β) Εισαγωγή πολλών δεικτών από την ανάγνωση μεταβλητών

Αφού λοιπόν ολοκληρώσαμε τη διαδικασία εμφάνισης του χάρτη και ενός σημείο ενδιαφέροντος ας δούμε πώς γίνεται να εμφανίσουμε πολλά σημεία ενδιαφέροντος καθώς διαβάζονται από ένα αρχείο δεδομένων. Η λογική είναι η ίδια με την εμφάνιση του ενός σημείου αλλά θα χρειαστούμε αντί για σταθερές συντεταγμένες, δύο μεταβλητές, που θα αλλάζουν κάθε φορά που αλλάζει και το σημείο ενδιαφέροντος. Αυτό σημαίνει ότι η προηγούμενη εντολή πρέπει να μπει σε μια δομή επανάληψης.

Τα δεδομένα τα διαβάζει ο κώδικας μέσα από ένα αρχείο που έχει διασυνδεδεμένη δομή όπως ένα XML/RDF, ένα OWL ακόμα και την ίδια την HTML. Παρακάτω βλέπουμε ένα παράδειγμα XML/RDF όπου γίνεται σαφές ότι αναφέρει

δύο αρχαία θέατρα, δίνοντας επίσης και πληροφορίες του συστήματος συντεταγμένων και τις συντεταγμένες στι οποίες αυτά βρίσκονται.

```
<gml:featureMember>
  <ogr:archaia_theatrafid="F1">
    <ogr:geometryProperty>
      <gml:PointsrsName="EPSG:2100">
        <gml:coordinates>181924.442195768002421,
          4386231.198737029917538</gml:coordinates>
      </gml:Point>
    </ogr:geometryProperty>
    <ogr:theatro>Αρχαίο Θέατρο των Γιτώνων
    </ogr:theatro>
    <ogr:y>4386231</ogr:y>
    <ogr:x>181924</ogr:x>
  </ogr:archaia_theatra>
</gml:featureMember>
<gml:featureMember>
  <ogr:archaia_theatra fid="F2">
    <ogr:geometryProperty>
      <gml:Point srsName="EPSG:2100">
        <gml:coordinates>315293.785974161990453,
          4116106.067466820124537</gml:coordinates>
      </gml:Point>
    </ogr:geometryProperty>
    <ogr:theatro>Ωδείο Μεσσήνης
    (Εκκλησιαστήριο)</ogr:theatro>
    <ogr:y>4116106</ogr:y>
    <ogr:x>315293</ogr:x>
  </ogr:archaia_theatra>
</gml:featureMember>
```

Για να διαβάσει το παραπάνω αρχείο ο χάρτης που κατασκευάστηκε, θα πρέπει να προσθέσουμε άλλο ένα μικρό κομμάτι κώδικα στο Javascript που θα κάνει parsing όπως συνηθίζουμε να λέμε (ανάλυση) των δεδομένων του αρχείου XML/RDF. Ο κώδικας είναι ο εξής:

```
xmlhttp=new XMLHttpRequest();
xmlhttp.open("GET", "aiolikos_charths.gml",false);
xmlhttp.send();
xmlDoc=xmlhttp.responseXML;
var place=xmlDoc.getElementsByTagName("gml:featureMember");
```

Ζητάει να δημιουργηθεί ένα αίτημα για ανάγνωση του XML/RDF αρχείου στην HTML της εφαρμογής μας. Περιμένει και λαμβάνει τις συντεταγμένες ως απάντηση. Ο κώδικας Javascript που είδαμε σε προηγούμενο κεφάλαιο μετατράπηκε σε δομή επανάληψης και τώρα είναι ο παρακάτω.

```

for (i=0; i<place.length; i++)
{
    var x =
place[i].getElementsByTagName("ogr:x")[0].childNodes[0].nodeValue;
    var y =
place[i].getElementsByTagName("ogr:y")[0].childNodes[0].nodeValue;
    var markersLayer = new OpenLayers.Layer.Markers("My Marker Layer");
    var iconSize = new OpenLayers.Size(20,20);
    var iconOffset = new OpenLayers.Pixel(-(iconSize.w/2), -iconSize.h);
    var q = new OpenLayers.LonLat(x,y).transform(new
OpenLayers.Projection("EPSG:2100"), new
OpenLayers.Projection("EPSG:900913"));
    var marker = new OpenLayers.Marker(q, new OpenLayers.Icon('map-
pointer.png',iconSize,iconOffset));
    markersLayer.setVisibility(true);
    markersLayer.addMarker(marker);
    map.addLayer(markersLayer);
}

```

Οι συντεταγμένες δηλαδή αντικαταστάθηκαν από τις μεταβλητές x και y. Ενδιαφέρον παρουσιάζει το γεγονός ότι ενώ δηλώσαμε ότι ο κώδικας θα ψάξει μέσα στο πεδίο "gml:featureMember", αργότερα ορίζουμε σε ποιο υποσύνολο του κόμβου θέλουμε ακριβώς να ψάξει και με ποιο όνομα αυτό δηλώνεται.

4.7 Parsing δεδομένων στο Wandora και ολοκλήρωση του διαδικτυακού τόπου

Ο HTML κώδικας του διακομιστή της εφαρμογής Wandora είναι έτσι κατασκευασμένος ώστε όταν θέσουμε ένα ερώτημα SPARQL, ο πίνακας που εμφανίζεται στην οθόνη μας, να έχει γραφτεί σε μιας μορφής δομημένης HTML. Τελικά εκμεταλλευόμενοι τα δομημένα αυτά δεδομένα μπορούμε να εξάγουμε τα αποτελέσματα σαν να ήταν σε κάποιο XML/RDF. Ακολουθώντας δηλαδή την ίδια λογική των διασυνδεδεμένων δεδομένων, προσθέτουμε στο Javascript το παρακάτω κομμάτι κώδικα.

```

var lonelements = document.getElementsByClassName("lon");
var latelements = document.getElementsByClassName("lat");
var nameslon = "";
var nameslat = "";
for (var i=0; i<features.length; i++) {
    nameslon = lonelements[i].innerHTML;
    nameslat = latelements[i].innerHTML;
}

```

```

        features[i] = new OpenLayers.Feature.Vector(
new OpenLayers.Geometry.Point(
nameslon, nameslat
).transform(new OpenLayers.Projection("EPSG:4326"), new
OpenLayers.Projection("EPSG:900913")), {
foo: val,
n : (val / (features.length))*100
}
);
}

```

Αυτό σημαίνει να πάρει όλα εκείνα τα στοιχεία τα οποία βρίσκει να έχουν ως όνομα κλάσης τα lon (longitude) και lat (latitude), να τα αντιστοιχίσει σε δύο μεταβλητές και για κάθε ένα στοιχείο να το περνάει στη μνήμη για εμφάνιση αργότερα του δείκτη. Η μεταβλητή foo χρησιμοποιείται για να εμφανίζεται κάτω από τον δείκτη μας το πλήθος των εγγραφών ή άλλων στατιστικών στοιχείων, ανάλογα με το τι θέλουμε να εμφανίσουμε κάθε φορά.

Σε αυτό ακριβώς το σημείο, θέλω να τονίσω τον περιορισμό που τίθεται λόγω της ονομασίας των δύο αυτών μεταβλητών lon και lat. Ο περιορισμός είναι ότι κατά τη σύνταξη του ερωτήματος SPARQL που θέλουμε, οι εντολές που θα μας επιστρέψουν τις συντεταγμένες πρέπει να μετονομασθούν κατά ανάγκη ?lon και ?lat αντίστοιχα για να διαβαστούν από τον κώδικα Javascript, αλλιώς δεν θα εμφανιστεί κανένας δείκτης στο χάρτη.

Το γραφικό στοιχείο του δείκτη άλλαξε, για να μας δίνει ένα φάσμα διαβάθμισης χρωμάτων ανάλογα με το ποσοστό επί τις εκατό που εμφανίζεται ένα στοιχείο στην ίδια χώρα. Μεταβλητή ελέγχου του ποσοστού είναι η n. Δηλαδή στο παράδειγμα SPARQL με τα συγκροτήματα, μια χώρα με 600 συγκροτήματα από τα σύνολο 1000 καταγεγραμμένα θα φαίνεται με πιο θερμό χρώμα από αυτήν που βρέθηκαν 20 συγκροτήματα. Η διαβάθμιση των χρωμάτων έγινε ανά δέκα τις εκατό αυξανόμενο ποσοστό. Παρακάτω δίνεται ο κώδικας για την εμφάνιση της διαβάθμισης των χρωμάτων και οι κανόνες που ακολουθούνται για την ορθή οπτικοποίησή τους.

```

var style = new OpenLayers.Style(
{
graphicWidth: 21,
graphicHeight: 25,
graphicYOffset: -28,
label: "${foo}"
},

```

Αφορά στο μέγεθος των δεικτών, πόσα pixels δηλαδή μέγεθος θα έχουν οι δείκτες ως προς μήκος και ύψος του δείκτη όπως επίσης και το τι θα γράφει η ετικέτα του

δείκτη. Στην πτυχιακή αυτή εργασία επέλεξα η ετικέτα να είναι το πλήθος των στοιχείων που βρέθηκε ανά χώρα και όχι το ποσοστό καθώς από αυτό θα ορίζεται μόνο το χρώμα. Με αυτόν τον τρόπο αποσπούμε δύο πληροφορίες στατιστικής σημασίας για την κάθε χώρα. Παρακάτω γράφουμε τους κανόνες που ακολουθεί το πρόγραμμα για την ορθή προβολή της διαβάθμισης των χρωμάτων των δεικτών.

Για κανόνα από 0 έως και 10 τις εκατό:

```
{
  rules: [
    new OpenLayers.Rule({
      filter: new OpenLayers.Filter.Comparison({
        type: OpenLayers.Filter.Comparison.LESS_THAN,
        property: "n",
        value: 10
      }),
      symbolizer: {
        externalGraphic: "${staticbase}marker-blue.png"
      }
    })
  ],
}
```

Για κανόνα 10 έως και 20 τις εκατό:

```
new OpenLayers.Rule({
  filter: new OpenLayers.Filter.Comparison({
    type: OpenLayers.Filter.Comparison.BETWEEN,
    property: "n",
    lowerBoundary: 10,
    upperBoundary: 20
  }),
  symbolizer: {
    externalGraphic: "${staticbase}marker-green.png"
  }
}),
```

Το ίδιο με παραπάνω για 20 έως και 90 τις εκατό. Για 90 με 100 τις εκατό όμως αλλάζει και είναι το παρακάτω:

```
new OpenLayers.Rule({
  elseFilter: true,
  symbolizer: {
    externalGraphic: "${staticbase}marker.png"
  }
})
```

Έτσι λοιπόν δημιουργήθηκε και η διαβάθμιση των χρωμάτων και η εμφάνιση του πλήθους των στοιχείων όπως φαίνεται στην παρακάτω εικόνα, θέτοντας στον κώδικα να διαβάσει random συντεταγμένες.



Σχήμα 4.9 : Παραγωγή τυχαίων συντεταγμένων και ετικετών στο χάρτη

Για το προηγούμενο Σχήμα 4.9 έχουμε τους παρακάτω κανόνες διαβάθμισης χρωμάτων:

Μπλε: 0 - 25 %

Πράσινο: 25 - 50 %

Κίτρινο: 50 - 75 %

Κόκκινο: 75 – 100%

Η κατασκευή του διαδικτυακού τόπου του Wandora τελειοποιήθηκε με την ένωση όλων των προαναφερθέντων τεχνολογιών σε ένα ενιαίο αρχείο HTML και δίνεται στο Παράρτημα I. Όμως ακόμα και στον έτοιμο κώδικα του διακομιστή της εφαρμογής υπήρξαν πολύ σημαντικές αλλαγές. Πιο συγκεκριμένα, προστέθηκαν element classNames και element ids έτσι ώστε να αναγνωρίζει το Javascript του χάρτη ποιά δεδομένα του κώδικα τελικά να αναλύει. Επίσης με ένα νέο Javascript γίνεται πλέον απόκρυψη του πίνακα των αποτελεσμάτων για να εμφανίζεται μόνο ο χάρτης χωρίς όμως την απώλεια δεδομένων.

Πριν από οποιαδήποτε τροποποίηση του κώδικα, για ένα μόνο ερώτημα SPARQL στο Wandora έχουμε σαν αποτέλεσμα έναν πίνακα με αποτελέσματα. Αν δημιουργήσουμε παραπάνω από ένα ερωτήματα, τότε αυτά εμφανίζονται ως instances, και πρέπει ο χρήστης να επιλέξει ποιά από όλα θέλει να εμφανιστεί.

Μετά την επεξεργασία του κώδικα, δίνεται η δυνατότητα για όσα ερωτήματα SPARQL θέλουμε, χωρίς περιορισμούς, τα οποία όμως δεν εμφανίζονται ως instances αλλά ως ένα μενού επιλογών στο πάνω αριστερά μέρος της ιστοσελίδας, πάνω ακριβώς από το χάρτη. Αυτό έγινε με την βοήθεια της γλώσσας προγραμματισμού CSS, πολύ ενδιαφέρον και με πολλές δυνατότητες εργαλείο, κατάλληλο για την δημιουργία γραφικών και μενού σε ιστοσελίδες. Θα το δούμε αναλυτικά παρακάτω.

Σε αυτό το σημείο να τονιστεί, ότι η εφαρμογή αυτή, όπως έχει κατασκευαστεί τώρα, μπορεί να δείξει μόνο στατιστικά στοιχεία και όχι τα ονόματα των στοιχείων που εμείς ψάχνουμε. Στο παράδειγμα με τα συγκροτήματα δηλαδή, το ερώτημα θα μας επιστρέψει στο παρασκήνιο και τα ονόματα των συγκροτημάτων, ο χάρτης όμως διαβάζει μόνο το πλήθος τους για την κάθε χώρα, βγάζοντας έτσι κάποιο στατιστικό συμπέρασμα. Ας δούμε τις επόμενες εικόνες πριν και μετά την επεξεργασία.

4.8 Κατασκευή του μενού της ιστοσελίδας και διορθώσεις

Σε αυτό το σημείο θα αναφερθούμε στις αισθητικές περισσότερο αλλαγές που χρειάστηκαν ώστε ο διαδικτυακός τόπος της εφαρμογής να είναι κατανοητός και φιλικός προς το χρήστη. Αρχικά για να γίνει αυτό έπρεπε να κατασκευαστεί ένα μενού επιλογών με τα ερωτήματα που κάθε φορά θέλει ο χρήστης να απεικονισθούν. Το μενού κατασκευάστηκε σε γλώσσα γραφικών CSS και η δομή της φαίνεται παρακάτω:

```
<style>
  nav ul ul {
    display: none;}
  nav ul li:hover > ul {
    display: block;}
  nav ul {
    background: #efefef;
    background: linear-gradient(top, #efefef 0%, #bbbbbb 100%);
    background: -moz-linear-gradient(top, #efefef 0%, #bbbbbb 100%);
    background: -webkit-linear-gradient(top, #efefef 0%, #bbbbbb 100%);
    box-shadow: 0px 0px 9px rgba(0,0,0,0.15);
    padding: 0 20px;
    border-radius: 10px;
    list-style: none;
    position: relative;
    display: inline-table;
  }
</style>
```

Τοποθετούμε το παραπάνω κομμάτι κώδικα στο <head> της ιστοσελίδας HTML. Οι πρώτες τέσσερις γραμμές του κώδικα αφορούν στην εμφάνιση και απόκρυψη του μενού όταν περνά το ποντίκι από πάνω, και όλες οι υπόλοιπες δίνουν τα χρώματα, τις σκιές και τη θέση του μενού που δημιουργήσαμε. Το υλοποιημένο μενού απεικονίζεται στο Σχήμα 4.10.

Οι διορθώσεις που έγιναν αφορούν περισσότερο στη δημιουργία ονομάτων των topics καθώς το Wandora από μόνο του δίνει ως ονομασία στα αποτελέσματα SPARQL Result Set και μια διεύθυνση. Επίσης επειδή όπως έχουμε πάλι αναφέρει τα δεδομένα και τα κείμενα στην Wikipedia τα γράφει ο οποιοσδήποτε, υπάρχουν περιπτώσεις σύγχυσης, όπου οι δείκτες σε μια περιοχή θα εμφανιστούν πάνω από μία φορές. Οπότε έγινε η απαραίτητη επεξεργασία και διόρθωση των αποτελεσμάτων για όσο το δυνατόν πιο σωστό αποτέλεσμα.



Σχήμα 4.10: Το υλοποιημένο μενού με τέσσερα διαφορετικά ερωτήματα.

4.9 Δημοσιοποίηση της εφαρμογής στο διαδίκτυο μέσω του Okeanos

Πριν την τελική δημοσιοποίηση των δεδομένων και γενικά της εφαρμογής, ας κάνουμε μια ανασκόπηση του τι ακριβώς δημιουργήσαμε. Αν τα προαναφερθέντα κομμάτια κώδικα Javascript, συγχωνευθούν σε ένα, και αυτό με τη σειρά του στην αρχική σελίδα HTML του διακομιστή του Wandora, τότε ο πίνακας του Σχήματος 4.6 εξαφανίζεται και στη θέση του εμφανίζεται ένας OSM χάρτης. Το αρχείο του Wandora με την αρχική σελίδα του διακομιστή βρίσκεται στο μονοπάτι C:\Users\User_Name_here\wandora\build\resources\server\topic\templates και ονομάζεται viewtopic.vhtml. Το επόμενο ακριβώς βήμα είναι να θέσουμε στην εφαρμογή Wandora οποιοδήποτε SPARQL ερώτημα θέλουμε να απεικονίσουμε, με τον περιορισμό όμως οι συντεταγμένες που θα επιστραφούν να έχουν τα ονόματα ?lon και ?lat. Αφού λοιπόν το Wandora εξάγει τα αποτελέσματά του από το διαδικτυακό τόπο της Dbpedia τότε μπορούμε να ανοίξουμε την αρχική σελίδα του διακομιστή του Wandora και να δούμε την οπτικοποίηση τους στον χάρτη.

Φυσικά, λόγω ελλειπών ή μη ορθών δεδομένων, κάποια topics χρειάζονται να υποστούν μια μικρή επεξεργασία για την όσο γίνεται πιο ορθή απεικόνισή τους. Συχνό είναι το φαινόμενο, σε κάθε χώρα να δηλώνονται δύο διαφορετικά ζευγάρια συντεταγμένων, δηλώνοντας την πρωτεύουσα και την συμπρωτεύουσα της κάθε μιας. Αποτέλεσμα είναι λοιπόν σε κάποιες περιπτώσεις να απεικονίζονται δύο δείκτες σε μια χώρα με το ίδιο ακριβώς πλήθος ως αριθμό ετικέτας τους.

Αφού ολοκληρωθεί και η επεξεργασία των δεδομένων, η εφαρμογή αυτή είναι έτοιμη για δημοσιοποίηση στον παγκόσμιο ιστό. Για να επιτευχθεί αυτό, έγινε χρήση της ελληνικής πρωτοβουλίας Okeanos, υπηρεσία παροχής διακομιστών σε φοιτητές και ερευνητές για ανάπτυξη νέων ερευνητικών και καινοτόμων εφαρμογών.

α) Η υπηρεσία ~okeanos

Η υπηρεσία αυτή παρέχεται από τον κεντρικό κόμβο της ΕΔΕΤ Α.Ε., υπεύθυνου φορέα παροχής διαδικτύου σε σχολεία και πανεπιστήμια, προς την ακαδημαϊκή και ερευνητική κοινότητα. Η χρήση του έχει ως στόχο την κάλυψη ακαδημαϊκών, εκπαιδευτικών και ερευνητικών αναγκών.

Οι υπολογιστικοί και αποθηκευτικοί πόροι του κόμβου μπορούν να χρησιμοποιηθούν ως διακομιστές και να αξιοποιηθούν είτε για μόνιμη φιλοξενία υπηρεσιών, είτε για πειραματικούς, ερευνητικούς και εκπαιδευτικούς σκοπούς από φοιτητές ή ερευνητικά προγράμματα.

Αυτό σημαίνει ότι ο κάθε ενδιαφερόμενος μπορεί να φτιάξει τον δικό του λογαριασμό στην υπηρεσία, που ταυτόχρονα του παρέχει έναν ή και περισσότερους εικονικούς υπολογιστές, οι οποίοι είναι πάντα συνδεδεμένοι στο διαδίκτυο, χωρίς να ανησυχεί για προβλήματα hardware, ή λογισμικού. Το hardware που μπορεί να επιλέξει δίνει το δικαίωμα για την κατασκευή ενός πολύ ισχυρού μηχανήματος για τα σημερινά δεδομένα, ενώ το λογισμικό μπορεί να είναι οποιοδήποτε κυκλοφορεί στον κόσμο αυτήν την στιγμή. Με την υπηρεσία Okeanos ο κάθε ενδιαφερόμενος είναι ένα κλικ μακριά από το δικό του εικονικό υπολογιστή και κατ'επέκταση από το δικό του εικονικό δίκτυο. Χρησιμοποιώντας οποιονδήποτε Web Browser υπάρχει άμεση σύνδεση με τον απομακρυσμένο αυτόν υπολογιστή τον οποίον μπορεί να τον διαχειριστή μέσω ενός Remote Desktop Connection. Αν φυσικά μετανιώσει ή τον βαρεθεί, μπορεί πολύ απλά να τον καταστρέψει και να δημιουργήσει έναν καινούριο με διαφορετικό λογισμικό ή με όποια άλλη παράμετρο τον βολεύει.

Υπάρχει επίσης η δυνατότητα να αποθηκεύει κάποιος τα αρχεία του σε απευθείας σύνδεση, να τα μοιράζεται με τους φίλους του και όλοι να έχουν ανά πάσα στιγμή πρόσβαση σε αυτά, από οπουδήποτε στον κόσμο μέσω των δικών τους εικονικών υπολογιστών. Παρέχει μια πλούσια εμπειρία της εξερεύνησης και ενασχόλησης με πολλά και διαφορετικά είδη του λογισμικού σε έναν υπολογιστή τον οποίον μπορούν να τον καταστρέψουν μετά την ολοκλήρωση της όποιας εργασίας έχουν αναλάβει. Καθιστά επίσης εύκολη την επαφή με διαφορετικές τεχνολογίες χωρίς να χρειάζεται να σπαταλήσει χρήματα ούτε για υλικό ούτε για λογισμικό. Είναι μια ζωντανή πλατφόρμα δοκιμών για τις σπουδές των φοιτητών, πάντα συνδεδεμένη με το διαδίκτυο.

Όσον αφορά στην αποθήκευση αρχείων, ο κάθε ενδιαφερόμενος χρήστης έχει τη δυνατότητα να γραφτεί στην υπηρεσία +rithos η οποία λειτουργεί σαν διαδικτυακός αποθηκευτικός χώρος παρέχοντας υπηρεσίες cloud με περίπου 50 Gb αποθηκευτικής μνήμης. Το λογισμικό που τροφοδοτεί την υπηρεσία αυτή ονομάζεται

"synnefo" και είναι υπηρεσία ανοιχτού κώδικα. Ο διαμοιρασμός αρχείων και υλικού γίνεται απλά, γρήγορα και με ασφάλεια μέσω ενός πολύ εύχρηστου και φιλικού προς τον χρήστη περιβάλλοντος εργασίας. Τέλος δίνει τη δυνατότητα συγχρονισμού των δεδομένων και των αρχείων με όλες τις συσκευές και τους υπολογιστές που είναι συνδεδεμένοι στο δίκτυο του okeanos.

Βέβαια ακόμα η υπηρεσία περιορίζεται για ερευνητική και επιστημονική χρήση από φοιτητές καθηγητές και ερευνητές, ενώ οι υπόλοιποι ενδιαφερόμενοι χρήστες μπορούν να την χρησιμοποιήσουν μόνο μετά από πρόσκληση στην υπηρεσία. Αυτό συμβαίνει γιατί ακόμα βρίσκεται σε άλφα έκδοση και δεν έχει τελειοποιηθεί ακόμη.

β) Οδηγίες δημιουργίας εικονικής μηχανής και εγκατάσταση της εφαρμογής

Ας ρίξουμε μια ματιά λοιπόν στην εγκατάσταση της εφαρμογής σε δημόσιο διακομιστή. Αρχικά δημιουργούμε έναν νέο λογαριασμό στην ιστοσελίδα okeanos.gnet.gr και αφού ενεργοποιηθεί, τότε δημιουργούμε έναν εικονικό υπολογιστή ο οποίος να έχει δυνατά χαρακτηριστικά για να μπορέσει να τρέξει χωρίς περιορισμούς και καθυστερήσεις το Wandora. Συνιστώμενα χαρακτηριστικά είναι 4Gb RAM, 4-core processor και 20Gb σκληρού δίσκου. Προσωπικά διάλεξα Windows Server 2012 για λειτουργικό σύστημα για σταθερότητα και ευκολία στη χρήση. Αφού λοιπόν δημιουργήσουμε τον εικονικό υπολογιστή, παρατηρούμε ένα παράθυρο που φαίνεται στο Σχήμα 4.11.



Σχήμα 4.11: Ο εικονικός υπολογιστής είναι ανοιχτός και έτοιμος για χρήση.

Επιλέγουμε τον υπολογιστή και αμέσως ένα παράθυρο ζητάει κώδικό πρόσβασης του Διαχειριστή του εικονικού αυτού υπολογιστή. Μόλις αυτός εισαχθεί, ανοίγει το Remote Desktop Connection των Windows του υπολογιστή μας και

χειριζόμαστε πλέον απομακρυσμένα και μέσω του διαδικτύου τον καινούριο μας υπολογιστή.

Τοποθετούμε όλα τα αρχεία τα οποία φτιάξαμε στο +rithos από τον υπολογιστή μας για να τα μεταφέρουμε στον εικονικό υπολογιστή. Κατεβάζουμε από το διαδικτυακό τόπο του Wandora την εφαρμογή, αλλά προσοχή. Για να λειτουργήσει, πρέπει να κατεβάσουμε την έκτη έκδοση της Java, δηλαδή την JRE 6.0 από το διαδικτυακό τόπο της Oracle. Μετά ανοίγουμε το Wandora και εκτελούμε τα ερωτήματα SPARQL που θέλουμε.

Αφού ολοκληρώσουμε τις παραπάνω διαδικασίες, αντιγράφουμε τα νέα αρχεία με τον κώδικα στο μονοπάτι του διακομιστή του Wandora και η εφαρμογή είναι έτοιμη να τρέξει τοπικά στον εικονικό υπολογιστή μας. Για να απενεργοποιήσουμε την τοπική σύνδεση και να καταφέρει να δώσει τα δεδομένα στο διαδίκτυο θα πρέπει να ακολουθήσουμε τα δύο παρακάτω βήματα.

Βήμα 1ο: Μέσα από την εφαρμογή Wandora, πηγαίνω στις επιλογές του Server από το μενού επιλογών. Το μονοπάτι είναι Server → Server settings. Αμέσως ανοίγει το παράθυρο το οποίο φαίνεται στο Σχήμα 4.12. Μετά απενεργοποιώ τα κουτάκια τα οποία αναγράφουν local only. Με ενδιαφέρει το πρώτο όπως απεικονίζεται και η κατηγορία Topic που βρίσκεται λίγο παρακάτω στο ίδιο παράθυρο.



Σχήμα 4.12:Server settings

Βήμα 2ο: Απενεργοποιώ το firewall των Windows Server του εικονικού υπολογιστή. Αυτό το κάνω γιατί για λόγους ασφαλείας του μηχανήματος η Microsoft έχει προνοήσει να προστατεύει τον υπολογιστή από εξωτερικούς χρήστες που θέλουν να συνδεθούν σε αυτόν. Εμείς όμως αυτό ακριβώς θέλουμε. Να συνδεθούν στην

εφαρμογή όσο το δυνατόν περισσότεροι. Η απενεργοποίηση του firewall γίνεται από τον πίνακα ελέγχου ακολουθώντας το μονοπάτι Control Panel\All Control Panel Items\Windows Firewall.

Όλα είναι έτοιμα και η εφαρμογή είναι έτοιμη για χρήση στο διαδίκτυο. Το όνομα της ιστοσελίδας που την φιλοξενεί μπορούμε να το βρούμε πολύ εύκολα ακολουθώντας τις παρακάτω οδηγίες. Πρώτον ανοίγουμε το Command Window της εικονικής μηχανής και πληκτρολογούμε ipconfig /all και μετά Enter. Τότε θα εμφανιστούν οι παρακάτω πληροφορίες.

```
Physical Address . . . . . : F0-7D-68-FD-A6-8D
DHCP Enabled. . . . . : Yes
Autoconfiguration Enabled . . . . : Yes
Link-local IPv6 Address . . . . . : fe80::83b:ca75:854f:2d75%11(Preferred)
IPv4 Address. . . . . : 192.168.2.7(Preferred)
Subnet Mask . . . . . : 255.255.255.0
Lease Obtained. . . . . : Σάββατο, 9 Φεβρουαρίου 2013 12:23:34 πμ
Lease Expires . . . . . : Δευτέρα, 11 Φεβρουαρίου 2013 12:23:35 πμ
Default Gateway . . . . . : 192.168.2.1
DHCP Server . . . . . : 192.168.2.1
DHCPv6 IAID . . . . . : 250641768
DHCPv6 Client DUID. . . . . : 00-01-00-01-18-1A-B1-D0-F0-7D-68-FD-A6-8D
```

Σχήμα 4.13:Command Window

Η διεύθυνση του διαδικτυακού τόπου της εφαρμογής μας είναι αυτή που αναγράφεται στο πεδίο IPv4 ακολουθούμενο από '/' και την πόρτα που μας έχει δώσει το Wandora και μετά '/topic' όπως φαίνεται στο Σχήμα 4.11. Άρα σύμφωνα με τα σχήματα 4.11 και 4.12 το site είναι το 192.168.2.7:8898/topic. Με αυτόν τον τρόπο, επιτυγχάνουμε την πρόσβαση στην εφαρμογή από οποιοδήποτε σημείο του πλανήτη μέσω του διαδικτύου.



Κεφάλαιο 5

Συμπεράσματα και μελλοντικές επεκτάσεις

5.1 Σύνοψη της πτυχιακής εργασίας

Σε αυτό το κεφάλαιο θα κάνουμε μια σύνοψη της πτυχιακής εργασίας και θα αναλύσουμε συμπεράσματα και τις μελλοντικές επεκτάσεις. Επίσης θα δούμε γιατί η πτυχιακή αυτή εργασία είναι καινοτόμα και γιατί διαφέρει από τις άλλες πρωτοβουλίες. Ας δούμε πάλι τι πραγματεύεται το κάθε κεφάλαιο ξεχωριστά.

Στην αρχή, στο πρώτο και δεύτερο κεφάλαιο της πτυχιακής εργασίας αναλύθηκε ο όρος του διαδικτύου και του Παγκόσμιου Ιστού γενικότερα μέσω μιας ιστορικής αναδρομής και τελικά αναλύεται πώς περνάμε από το διαδίκτυο των εγγράφων και της απλής αναζήτησης πληροφορίας, στο διαδίκτυο των δεδομένων όπου η ανταλλαγή απόψεων, αρχείων και πολυμέσων καλπάζει, οι άνθρωποι όσο μακριά και να βρίσκονται επικοινωνούν, γνωρίζονται, διασκεδάζουν, ψυχαγωγούνται. Αργότερα αναλύεται το τελικό στάδιο στο οποίο έχει περιέλθει το διαδίκτυο, τον σημασιολογικό Ιστό δηλαδή. Έτσι αναλύονται όροι όπως σημασιολογία, ανοιχτά δεδομένα, masups, topic maps καθώς και νέες τεχνολογίες που χρησιμοποιούνται. Στα κεφάλαια αυτά ο όρος “Ανοιχτά Δεδομένα” δεν περνά απαρατήρητος καθώς αποτελεί μια πολύ σημαντική παράμετρο για την ελεύθερη και δωρεάν γνώση, προσβάσιμη από όλον τον κόσμο. Πρωτοβουλίες όπως της Wikipedia που είναι η μεγαλύτερη εγκυκλοπαίδεια στο διαδίκτυο, να παρέχει δωρεάν και ανοιχτά όλα τα δεδομένα της, έχουν πάρει τον κόσμο με το μέρος τους και όλο και περισσότεροι βοηθούν εθελοντικά σε αυτές. Έτσι η έρευνα και ανακάλυψη της γνώσης γίνεται πιο εύκολα προσβάσιμη από ποτέ.

Στο τρίτο κεφάλαιο της εργασίας αυτής, είδαμε πώς με την βοήθεια διάφορων εργαλείων διαχείρισης των ανοιχτών δεδομένων μπορούμε να επεξεργασθούμε, και να δημοσιοποιήσουμε σε άλλους διαδικτυακούς τόπους τα διασυνδεδεμένα δεδομένα. Είδαμε μερικές με δικό τους σκοπό και έργο, και στην συνέχεια αναλύθηκαν όροι όπως τα topic maps, πολύ σημαντικού και ισχυρού εργαλείου οπτικοποίησης της διασυνδεδεμένης γνώσης. Μέσα από 211 διαφορετικά εργαλεία ξεχωρίσαμε για την εφαρμογή μας το Wandora και είδαμε μια πολύ γρήγορη περιήγηση στο περιβάλλον του κατανοώντας για μια ακόμη φορά πόσο εύκολη μπορεί να είναι η διαδικασία της δημιουργίας, επεξεργασίας και δημοσιοποίησης των δεδομένων. Τέλος αναλύονται όροι όπως topic, association και occurrence, βασικά δομικά στοιχεία των ανοιχτών διασυνδεδεμένων δεδομένων και με την υλοποίηση ενός παραδείγματος καταλήγουμε σε συμπεράσματα.

Το τέταρτο θεωρώ ότι είναι το πιο σημαντικό κεφάλαιο καθώς σε αυτό αναλύσαμε όλη τη δημιουργία της εφαρμογής και επεξηγήθηκαν παραδείγματα κώδικα και σύνταξης ερωτημάτων SPARQL. Είδαμε πώς δημιουργείται ένας γεωγραφικός χάρτης ο οποίος αποδίδει με διαβαθμίσεις χρωμάτων ανά χώρα ή όποιο άλλο στατιστικό στοιχείο του έχουμε ζητήσει. Μετά την υλοποίηση και τα αποτελέσματα βλέπουμε το πώς έγινε η διάθεσή της υλοποιημένης εφαρμογής στο okeanos.grnet.gr, αναφερθήκαμε στην πρωτότυπη αυτή πρωτοβουλία που αποδεικνύει ότι η χώρα μας παρά τις δύσκολες συνθήκες μπορεί να πάει τεχνολογικά πολύ μπροστά. Τέλος δόθηκαν οδηγίες για εγκατάσταση και έναρξη λειτουργίας της εφαρμογής από έναν απομακρυσμένο διακομιστή προς το διαδίκτυο.

Έτσι φτάνουμε στο 5ο και τελευταίο κεφάλαιο της εργασίας, όπου αναλύονται τα συμπεράσματα και προτεινόμενες μελλοντικές επεκτάσεις της εφαρμογής αυτής καθώς επίσης και κίνητρα για ανάπτυξη περισσότερων νέων παρόμοιων εφαρμογών.

Να σημειωθεί ότι ο κώδικας και άλλα χρήσιμα στοιχεία βρίσκονται στα Παραρτήματα I και II.

5.2 Συμπεράσματα

Μετά την προσθήκη όλων των προγραμμάτων, που αναλύσαμε στο προηγούμενο κεφάλαιο, σε ένα μόνο κομμάτι κώδικα σε σελίδα γραμμένη σε HTML, ολοκληρώνεται η υλοποίηση της διαδικτυακής εφαρμογής απεικόνισης στατιστικών δεδομένων σε γεωγραφικούς χάρτες. Η τελική εφαρμογή αποτελείται πλέον από τον OpenStreetMap χάρτη στη θέση που πριν ήταν ο πίνακας με τα αποτελέσματα, και το μενού των επιλογών στο πάνω δεξιά μέρος της σελίδας. Στο κάτω δεξιά μέρος της σελίδας δίνεται ο χάρακας απεικόνισης των διαβαθμίσεων των χρωμάτων των δεικτών.

Σε αυτό το σημείο πρέπει να γίνει μια διευκρίνιση. Τα ποσοστά στο χάρακα δεν έχουν αποδοθεί ισοσκελώς και αυτό γιατί σπάνια όταν το πλήθος των αποτελεσμάτων είναι πολύ μεγάλο, φαίνονται διαφορές σε κάθε χώρα πάνω από 10%. Με περισσότερα χρώματα στις απεικονίσεις μικρών ποσοστών, βγάζουμε περισσότερα και πιο ομοιόμορφα συμπεράσματα στο χάρτη (Σχήμα 5.1).

Στο αριστερό κάτω μέρος της εφαρμογής δίνεται το ολικό ποσό των topics που έχουν βρεθεί στο διαδίκτυο και σύμφωνα πάντα με την Wikipedia και την Dbpedia. Πρέπει να τονιστεί ότι σε καμία περίπτωση τα αποτελέσματα που βγαίνουν είναι 100% ακριβή, και αυτό οφείλεται σε δύο κύριους παράγοντες. Ο πρώτος είναι ότι το Wandora ψάχνει μόνο στις διασυνδεδεμένες βάσεις δεδομένων της Wikipedia οπότε εγγραφές που δεν έχουν γίνει εκεί, αλλά δεν μπορούν να εμφανιστούν και ο δεύτερος παράγοντας είναι ότι υπάρχουν δυστυχώς ανακρίβειες και παραλείψεις. Όπως έχουμε αναφέρει ξανά, τα δεδομένα και τα κείμενα που καταγράφονται στην Wikipedia, ελέγχονται μόνο ως προς την ορθότητά τους και για τυχόν ψευδή στοιχεία, και όχι για παραλείψεις ή ανακρίβειες. Στο πάνω δεξιά μέρος της εφαρμογής μπορεί κάποιος να επιλέξει το σύνδεσμο “Back to Results” για να επιστρέψει στην αρχική σελίδα και να προβάλλει ένα από τα υπόλοιπα ερωτήματα της λίστας του μενού.

Για την προβολή της εφαρμογής στο διαδίκτυο έγινε η χρήση ενός δημόσιου διακομιστή από τον κόμβο της ΕΔΕΤ Α.Ε. Τελικά η εφαρμογή απεικόνισης στατιστικών αποτελεσμάτων σε γεωγραφικούς χάρτες με αποκλειστική χρήση ανοιχτών διασυνδεδεμένων δεδομένων είναι πλέον γεγονός. Και διαφέρει από τις άλλες εφαρμογές απεικόνισης σε χάρτες γιατί αντλεί ζωντανά, και από το διαδίκτυο δεδομένα χωρίς την εξάρτησή της από εξωτερικά XML/RDF αρχεία. Ας ρίξουμε μια ματιά στο παρακάτω Σχήμα 5.2 και στην τελική υλοποίηση της εφαρμογής, στο πως δηλαδή φαίνεται στο διαδίκτυο και αργότερα θα αναλύσουμε γιατί είναι διαφορετική από τις άλλες εφαρμογές.



Σχήμα 5.1: Οι διαβαθμίσεις των χρωμάτων και τα ποσοστά %

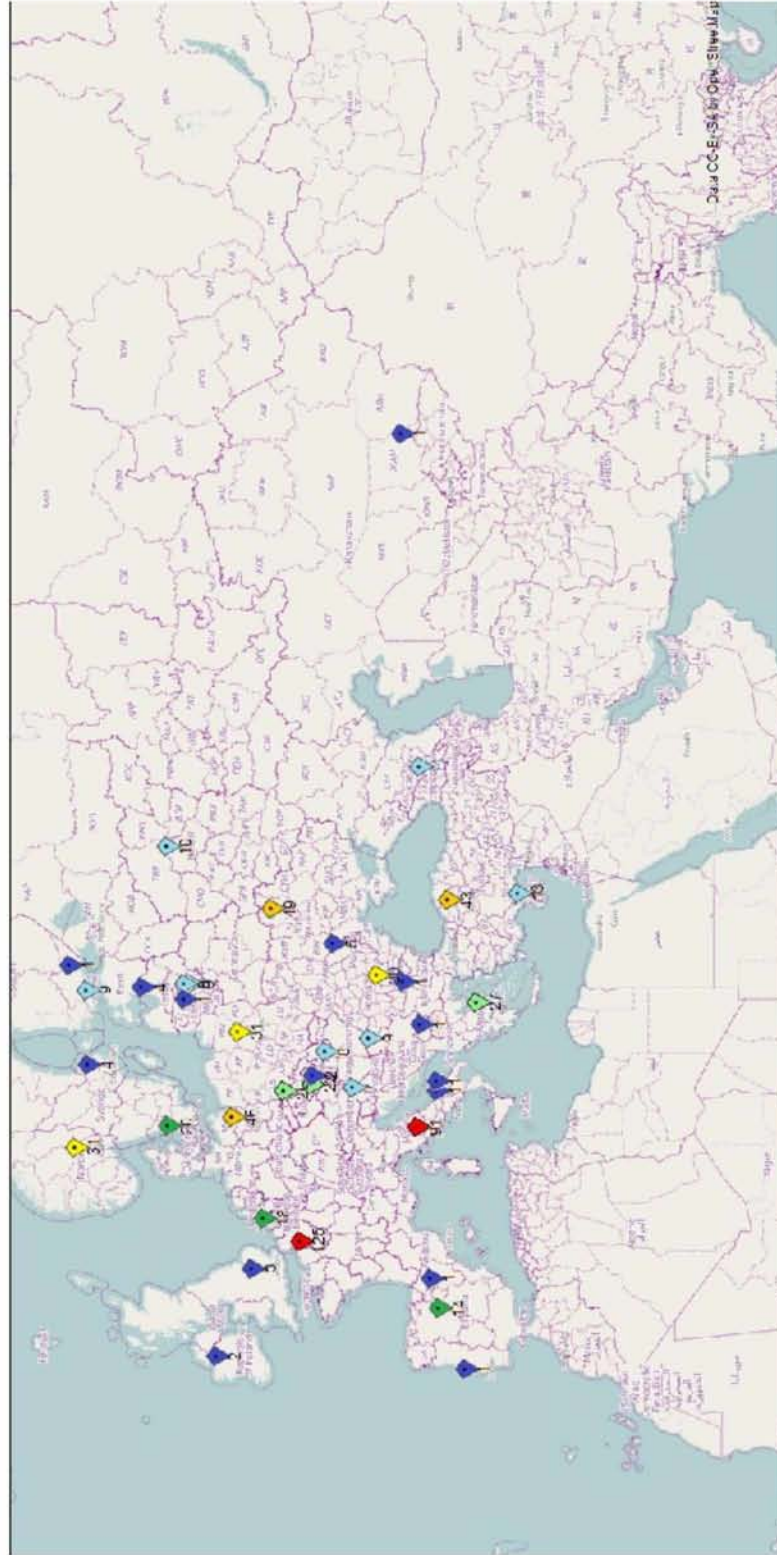
5.3 Η διαφορά της εφαρμογής από τις άλλες πρωτοβουλίες

Όπως προαναφέρθηκε υπάρχουν πολλές εφαρμογές απόδοσης ανοιχτών δεδομένων σε γεωγραφικούς χάρτες, παρόμοιες με την παρούσα εφαρμογή της πτυχιακής εργασίας. Όμως εκείνες οι εφαρμογές λειτουργούν σε διαφορετική λογική ανοιχτών δεδομένων. Η λογική τους βασίζεται στην απεικόνιση δεικτών, σύμφωνα με την καταγραφή τους σε αρχεία διασυνδεδεμένων δεδομένων τύπου XML/RDF. Στην περίπτωση των χαρτών βέβαια χρησιμοποιούνται για λόγους ευκολίας και ταχύτητας τύποι όπως KML ή GML που περιέχουν συγκεκριμένη δομή απόδοσης γεωγραφικών δεδομένων και διαβάζονται από έναν χάρτη πολύ πιο εύκολα.

Ας πούμε ότι ένα τέτοιο GML αρχείο περιέχει τα ονόματα Νοσοκομείων και τις γεωγραφικές συντεταγμένες τους. Αν κάποια στιγμή ανοίξει ένα καινούριο νοσοκομείο, ή γίνει κάποια μεταφορά ή αλλαγή ονόματος, τότε θα πρέπει να παρέμβει κάποιος να τροποποιήσει και το GML αρχείο με τα νέα δεδομένα. Ναι μεν δηλαδή τα δεδομένα είναι ανοιχτά προς όλους, αλλά δεν είναι απόλυτα διασυνδεδεμένα στο εύρος του διαδικτύου.

Στην εφαρμογή που αναπτύχθηκε στο πλαίσιο της πτυχιακής εργασίας ακολουθείται διαφορετική λογική. Μέσω του περιβάλλοντος Wandora εκτελούνται live ερωτήματα στο διαδίκτυο και επιστρέφονται αποτελέσματα με όλες τις αλλαγές που έχουν γίνει. Αν δηλαδή για το ίδιο παράδειγμα, αλλάξει κάτι που αφορά στα νοσοκομεία μια χώρας, τότε σίγουρα θα γραφτεί στα πρακτικά και κατ'επέκταση στο διαδικτυακό τόπο του εκάστοτε υπεύθυνου φορέα. Με το που καταγραφθεί κάτι, αμέσως ενημερώνεται και ο χάρτης, χωρίς να εξαρτάται από την παρουσία και διενέργεια από τρίτο πρόσωπο. Αυτό αποτελεί κάτι καινοτόμο για την σημερινή λογική που λειτουργεί το διαδίκτυο και αυτό είναι που μας κάνει να διαφέρουμε από τις υπόλοιπες πρωτοβουλίες.

Ανασχεδίαση of bymap



Topic: Total found 691

Wandora - Helmas application developed by George Papadopoulos
 University of Central Greece
 Department of Computer Science and Informatics



Σχήμα 5.2 : Ο χάρτης μέσω του Wandora μετά την επεξεργασία

5.4 Μελλοντικές επεκτάσεις

Τέλος θα ήθελα να αποτυπώσω κάποιες ιδέες για την μελλοντική επέκταση και αναβάθμιση της εφαρμογής που αναπτύχθηκε στο πλαίσιο της πτυχιακής εργασίας. Όπως κάποιος μπορεί να καταλάβει, οι επιλογές που παρέχουν τα ανοιχτά διασυνδεδεμένα δεδομένα είναι πάρα πολλές και ανοίγουν νέους ορίζοντες στην δημιουργία νέων ισχυρών και δυναμικών εφαρμογών.

Οπότε σε αυτό το σημείο, θα ήθελα να προτείνω μερικές επεκτάσεις για το μέλλον. Πρώτον, θα μπορούσε ο δείκτης της κάθε χώρας να είναι δυναμικός και να μπορεί να επιλεγεί από τον χρήστη. Τότε θα μπορούσε να ανοίγει σε νέο παράθυρο ο πίνακας με τα ονόματα των topics που αντιστοιχούν στην συγκεκριμένη χώρα. Για παράδειγμα, αν γινόταν να πατηθεί ο πράσινος δείκτης του Σχήματος 5.2 που δείχνει στην Ελλάδα, τότε να άνοιγε μια νέα καρτέλα με τα ονόματα όλων των Εκπαιδευτικών Ιδρυμάτων της χώρας.

Μία δεύτερη ιδέα είναι η δυνατότητα από τον χρήστη SPARQL ερωτήματος κατευθείαν στο διαδικτυακό τόπο της εφαρμογής σε μια επιλογή “Άλλο” στη λίστα του μενού. Έτσι θα δίνεται η δυνατότητα ο καθένας που δεν βρίσκει κάποιο ερώτημα που να τον εκφράζει, να αναπαράγει το δικό του σε μια SPARQL φόρμα, που ίσως και να αποθηκεύεται στην λίστα με τα υπόλοιπα ερωτήματα.

Μια τρίτη και δύσκολη αλλά σίγουρα πολύ χρήσιμη ιδέα γενικότερα για τα ανοιχτά διασυνδεδεμένα δεδομένα θα ήταν να μπορεί να γίνει διόρθωση των ανακρίβειών και παραλείψεων κατευθείαν στους διαδικτυακούς τόπους από τους οποίους λήφθηκαν τα δεδομένα. Για παράδειγμα, όταν μια χώρα δηλώνεται ως subject χωρίς όμως την διασύνδεσή της με την αντίστοιχη σελίδα, μετά την επεξεργασία από τον χρήστη να αποστέλλεται suggestion στην Wikipedia ή την Dbpedia για άμεση διόρθωση.

Βιβλιογραφικές Αναφορές

Wandora

<http://www.wandora.org/www/>

<http://www.wandora.org/wiki/Quickstart>

<http://www.wandora.org/wiki/Documentation>

[http://www.wandora.org/wiki/Documentation#Topic Maps](http://www.wandora.org/wiki/Documentation#Topic_Maps)

<http://www.wandora.org/forum/>

OpenLayers & Openstreetmap

<http://www.openlayers.org/>

<http://www.openstreetmap.org/>

<http://www.openlayers.org/dev/examples/example-list.html>

[http://wiki.openstreetmap.org/wiki/Main Page](http://wiki.openstreetmap.org/wiki/Main_Page)

Dbpedia

<http://dbpedia.org/About>

<http://dbpedia.org/Ontology>

<http://dbpedia.org/sparql>

<http://www.w3.org/TR/rdf-sparql-query/#groupExamples>

Διάφοροι σύνδεσμοι

<http://www.ns-tech.co.uk/blog/2008/08/using-openstreetmap-overview-part-1-obtaining-mapping-data/>

ΠΑΡΑΡΤΗΜΑ Ι

Κώδικας των χαρτών στον server του Wandora σε Javascript.

```
<div id="map"></div>

<script type="text/javascript">

//Διάβασμα των συντεταγμένων από το parsing των
//δεδομένων

var lonelements = document.getElementsByClassName("lon");
var latelements = document.getElementsByClassName("lat");
var valuelements = document.getElementsByClassName("value");
var nameslon = "";
var nameslat = "";
var values = "";

//Δημιουργία του χάρτη
var map = new OpenLayers.Map("map");
var wms = new OpenLayers.Layer.OSM("OpenStreetMap");

/*Parsing των δεδομένων μέσα από την VHTML, μετατροπή των συντεταγμένων σε
σφαιρικό σύστημα
συντεταγμένων και εισαγωγή δύο μεταβλητών, μια για το
πλήθος και μια για το ποσοστό.*/

var features = new Array(lonelements.length);
  for (var i=0; i<features.length; i++)
  {
      nameslon = lonelements[i].innerHTML;
      nameslat = latelements[i].innerHTML;
      val = valuelements[i].innerHTML;
      features[i] = new OpenLayers.Feature.Vector(
        new OpenLayers.Geometry.Point(
          nameslon, nameslat

          ).transform(new OpenLayers.Projection("EPSG:4326"),    new
OpenLayers.Projection("EPSG:900913")), {

          foo: val,
          n : (val / (features.length))*100

        }
      );
  }
}
```

*/*Εδώ δημιουργείται το style των δεικτών. Ποιούς κανόνες εμφάνισης ακολουθούν και ποιο θα είναι το label κάτω από κάθε έναν δείκτη. Οι κανόνες ακολουθούν τις εντολές LESS THAN και BETWEEN για την κατανομή του ποσοστού*/*

```
var style = new OpenLayers.Style(  
  {  
    graphicWidth: 21,  
    graphicHeight: 25,  
    graphicYOffset: -28,  
    label: "${foo}"  
  },  
  
  {  
    rules: [  
      new OpenLayers.Rule({  
  
        filter: new OpenLayers.Filter.Comparison({  
          type:  
OpenLayers.Filter.Comparison.LESS_THAN,  
          property: "n",  
          value: 10  
        }),  
  
        symbolizer: {  
          externalGraphic: "${staticbase}marker-blue.png"  
        }  
      }),  
      new OpenLayers.Rule({  
        filter: new OpenLayers.Filter.Comparison({  
          type: OpenLayers.Filter.Comparison.BETWEEN,  
          property: "n",  
          lowerBoundary: 10,  
          upperBoundary: 20  
        }),  
        symbolizer: {  
          externalGraphic: "${staticbase}marker-green.png"  
        }  
      }),  
      new OpenLayers.Rule({  
        filter: new OpenLayers.Filter.Comparison({  
          type: OpenLayers.Filter.Comparison.BETWEEN,  
          property: "n",  
          lowerBoundary: 20,  
          upperBoundary: 30  
        }),  
        symbolizer: {  
          externalGraphic: "${staticbase}marker-gold.png"  
        }  
      }  
    ]  
  }  
);
```

```

    }},
        new OpenLayers.Rule({
        filter: new OpenLayers.Filter.Comparison({
            type: OpenLayers.Filter.Comparison.BETWEEN,
            property: "n",
            lowerBoundary: 30,
            upperBoundary: 40
        }),
        symbolizer: {
            externalGraphic: "${staticbase}marker-gold.png"
        }
    }},
        new OpenLayers.Rule({
        filter: new OpenLayers.Filter.Comparison({
            type: OpenLayers.Filter.Comparison.BETWEEN,
            property: "n",
            lowerBoundary: 40,
            upperBoundary: 50
        }),
        symbolizer: {
            externalGraphic: "${staticbase}marker-gold.png"
        }
    }},
        new OpenLayers.Rule({
        filter: new OpenLayers.Filter.Comparison({
            type: OpenLayers.Filter.Comparison.BETWEEN,
            property: "n",
            lowerBoundary: 50,
            upperBoundary: 60
        }),
        symbolizer: {
            externalGraphic: "${staticbase}marker-gold.png"
        }
    }},
        new OpenLayers.Rule({
        filter: new OpenLayers.Filter.Comparison({
            type: OpenLayers.Filter.Comparison.BETWEEN,
            property: "n",
            lowerBoundary: 60,
            upperBoundary: 70
        }),
        symbolizer: {
            externalGraphic: "${staticbase}marker-gold.png"
        }
    }},
        new OpenLayers.Rule({
        filter: new OpenLayers.Filter.Comparison({

```

```

        type: OpenLayers.Filter.Comparison.BETWEEN,
        property: "n",
        lowerBoundary: 70,
        upperBoundary: 80
    }},
    symbolizer: {
        externalGraphic: "${staticbase}marker-gold.png"
    }
    }},
    new OpenLayers.Rule({
    filter: new OpenLayers.Filter.Comparison({
    type: OpenLayers.Filter.Comparison.BETWEEN,
    property: "n",
    lowerBoundary: 80,
    upperBoundary: 90
    }},
    symbolizer: {
        externalGraphic: "${staticbase}marker-gold.png"
    }
    }},
    new OpenLayers.Rule({
    elseFilter: true,
    symbolizer: {
        externalGraphic: "${staticbase}marker.png"
    }
    }
    }
    }
    );

```

```

var vector = new OpenLayers.Layer.Vector("Vector Layer", {
    styleMap: new OpenLayers.StyleMap(style)
});
vector.addFeatures(features);

```

```

map.addLayers([wms, vector]);
map.setCenter(new OpenLayers.LonLat(0, 0), 1);

```

```

</script>

```

Ερωτήματα SPARQL που χρησιμοποιήθηκαν στην ανάπτυξη της εφαρμογής.

```
SELECT (COUNT(DISTINCT ?band) AS ?value) ?origin ?cap ?lat ?lon
WHERE
{
  ?about dbpedia-owl:genre <http://dbpedia.org/resource/Hard_rock>;
  dbpedia-owl:associatedMusicalArtist ?band.
  ?band dbpprop:origin ?origin.
  ?origin dbpedia-owl:capital ?cap;
  geo:lat ?lat;
  geo:long ?lon.
}
```

Βρες το πλήθος των συγκροτημάτων που ανήκουν στην κατηγορία Hard_rock ανά χώρα στον κόσμο.

```
SELECT (COUNT(DISTINCT ?band) AS ?value) ?origin ?cap ?lat ?lon
WHERE
{
  ?about dbpedia-owl:genre <http://dbpedia.org/resource/Jazz>;
  dbpedia-owl:associatedMusicalArtist ?band.
  ?band dbpprop:origin ?origin.
  ?origin dbpedia-owl:capital ?cap;
  geo:lat ?lat;
  geo:long ?lon.
}
```

Βρες το πλήθος των συγκροτημάτων που ανήκουν στην κατηγορία Jazz ανά χώρα στον κόσμο.

```
SELECT DISTINCT (COUNT (DISTINCT ?unis )AS ?value) ?uni ?country ?lat ?lon
WHERE
{
  ?about skos:broader
  <http://dbpedia.org/resource/Category:Universities_and_colleges_in_Europe>.
  ?uni skos:broader ?about.
  ?unis dcterms:subject ?uni.
  ?unis dbpedia-owl:country ?country.
  ?country geo:lat ?lat.
  ?country geo:long ?lon.
}
```

Βρες το πλήθος των πανεπιστημιακών ιδρυμάτων και κολλεγίων ανά χώρα που ανήκει στην Ευρωπαϊκή Ένωση.


```

SELECT (COUNT(DISTINCT ?about) AS ?value) ?loc ?lat ?lon
WHERE
{
?about dbpedia-owl:product <http://dbpedia.org/resource/Automobile>;
dbpedia-owl:location ?loc.
?loc geo:lat ?lat.
?loc geo:long ?lon.
}

```

Βρες το πλήθος των βιομηχανιών αυτοκινήτων ανά χώρα στον κόσμο.

```

SELECT ?lon ?lat (COUNT (DISTINCT ?about) AS ?value)
WHERE
{
?about dcterms:subject <http://dbpedia.org/resource/Category:Greek_temples>;
geo:lat ?lat;
geo:long ?lon.
}

```

Βρες τους αρχαιοελληνικούς ναούς και εμφάνισέ τους στο χάρτη.

```

SELECT ?lon ?lat (COUNT (DISTINCT ?about) AS ?value)
WHERE
{
?about dcterms:subject
<http://dbpedia.org/resource/Category:Temples_in_ancient_Athens>;
geo:lat ?lat;
geo:long ?lon.
}

```

Βρες τους αρχαίους ναούς που βρίσκονται στην Αθήνα και εμφάνισέ τους στο χάρτη.

```

SELECT ?lat ?lon (COUNT (DISTINCT ?about) AS ?value)
WHERE
{
?about rdf:type <http://dbpedia.org/class/yago/ArchaeologicalMuseumsInGreece>;
geo:lat ?lat;
geo:long ?lon.
}

```

Βρες τα αρχαιολογικά μουσεία στην Ελλάδα και εμφάνισέ τα στο χάρτη.

ΠΑΡΑΡΤΗΜΑ ΙΙ

Επεξήγηση του κώδικα javascript

```
<div id="map"></div>
```

Αφορά στην οπτικοποίηση του χάρτη πάνω στην ιστοσελίδα. Όπου τοποθετηθεί το div εκεί θα εμφανιστεί ο χάρτης.

```
var lonelements = document.getElementsByClassName("lon");  
var latelements = document.getElementsByClassName("lat");  
var valuelements = document.getElementsByClassName("value");
```

Η εντολή `document.getElementsByClassName()`; επιλέγει ένα στοιχείο από τον πηγαίο κώδικα της ιστοσελίδας μας με όνομα αυτό που θα δηλώσουμε μέσα στην παρένθεση.

ΠΡΟΣΟΧΗ! Αυτό σημαίνει άρα ότι η μόνη αλλαγή που χρειάζεται στο γνήσιο vhtml του server, είναι να τοποθετήσουμε στα topics που μας ενδιαφέρουν ή που παίρνουν τα στοιχεία που θέλουμε να αναλύσουμε (parsing), το όνομα των κλάσεων που θέλουμε να χρησιμοποιήσουμε.

Π.χ. `<topic></topic>` θα το αλλάξουμε σε `<topic class="name"></topic>`

```
nameslon = lonelements[i].innerHTML;  
nameslat = latelements[i].innerHTML;  
val = valuelements[i].innerHTML;
```

Αφορά στο διάβασμα εσωτερικά του html των μεταβλητών lonelements, latelements και τοποθέτηση σε έναν πίνακα.

```
features[i] = new OpenLayers.Feature.Vector(  
    new OpenLayers.Geometry.Point(  
        nameslon, nameslat)
```

Το features αφορά στα στοιχεία ξεχωριστά που παίρνουμε μέσα από το HTML με όνομα nameslon και nameslat από τον πίνακα.

```
        .transform(new OpenLayers.Projection("EPSG:4326"), new  
        OpenLayers.Projection("EPSG:900913")),  
    Μετασχηματισμός των συντεταγμένων από EPSG:4326 σε EPSG:900913  
    {  
    foo: val,  
    n : (val / (features.length))*100  
    }
```

Και απεικόνιση (projection) των μεταβλητών foo και n που θα δούμε χρησιμοποιηθούν παρακάτω στην απεικόνιση του δείκτη.

Επεξήγηση του κώδικα CSS

Η μεταβλητή `style` είναι ένα είδος CSS γραμμένο εσωτερικά της `javascript`. Αφορά στην απεικόνιση των δεικτών και ιορίζει το στυλ το οποίο θέλουμε αυτές να εμφανίζονται.

```
graphicWidth: 21,  
graphicHeight: 25,  
graphicYOffset: -28,  
label: "{$foo}"
```

Οριοθέτηση σε pixel του δείκτη και πινακίδα `label` της `foo` μεταβλητής.

```
new OpenLayers.Rule({  
  filter: new OpenLayers.Filter.Comparison({  
    type: OpenLayers.Filter.Comparison.BETWEEN,  
    property: "n",  
    lowerBoundary: 10,  
    upperBoundary: 20
```

Κανόνες εισαγωγής του κατάλληλου χρώματος δείκτη και αριθμού της πινακίδας ανάλογα με το ποσοστό επί τις εκατό. Στο συγκεκριμένο παράδειγμα αφορά για μεταξύ 10 και 20 τις εκατό τι θα εμφανιστεί.

ΛΕΞΙΚΟ ΟΡΩΝ

Παρακάτω αναφέρονται οι σημαντικότεροι όροι σε αλφαβητική σειρά:

Heatmap: Γεωγραφικός χάρτης απεικόνισης στατιστικών δεδομένων σε κλίμακα διαβάθμισης χρωμάτων.

Mashup: Είναι μια εφαρμογή που μπορεί και συνδυάζει και να ομαδοποιεί δεδομένα με σκοπό να δημιουργήσει νέες υπηρεσίες. Συνοπτικά δηλαδή στην περίπτωση του διαδικτύου ο όρος mashup σημαίνει “μάζεμα” των ανοιχτών δεδομένων.

Okeanos: Είναι υπηρεσία παροχής διακομιστών σε φοιτητές και ερευνητές για ανάπτυξη νέων ερευνητικών και καινοτόμων εφαρμογών. Η υπηρεσία αυτή παρέχεται από τον κεντρικό κόμβο της ΕΔΕΤ Α.Ε., υπεύθυνου φορέα παροχής διαδικτύου σε σχολεία και πανεπιστήμια, προς την ακαδημαϊκή και ερευνητική κοινότητα. Η χρήση του έχει ως στόχο την κάλυψη ακαδημαϊκών, εκπαιδευτικών και ερευνητικών αναγκών.

OpenStreetMap (OSM): Είναι ένας από τους χάρτες που παρέχουν τα OpenLayers, είναι χάρτης ανοιχτού κώδικα και επεξεργάσιμος από τους χρήστες. Τα δεδομένα που καταγράφονταν από τα συστήματα GPS καθώς και πολλά ανοιχτά γεωγραφικά δεδομένα όπως αεροφωτογραφίες, συντεταγμένες και σύνορα πόλεων καθώς και πολλά στατιστικά δεδομένα δημιούργησαν τους χάρτες αυτούς. Η ιστοσελίδα του υποστηρίζεται από τον OpenStreetMap Foundation.

Semantics: Είναι η διευκρίνιση των όρων που περιγράφουν κάτι όταν αυτοί έχουν κοινή σημασία. Στον άνθρωπο οι λέξεις origin, source, principle, provenance έχουν κοινή σημασία. Με την ανάπτυξη σημασιολογικών τεχνολογιών, ο υπολογιστής μπορεί να κατανοήσει την κοινή αυτή σημασία.

Semantic Web: Σημασιολογική τεχνολογία για το διαδίκτυο. Ορίζεται ένα “δίκτυο δεδομένων το οποίο μπορούν να το επεξεργαστούν οι υπολογιστές άμεσα ή έμμεσα”. Αυτό σημαίνει δηλαδή ότι αν δοθεί σαφή έννοια για τον κάθε πόρο ή θέμα που πραγματεύεται μια ιστοσελίδα, τότε κατά κάποιο τρόπο ο υπολογιστής καταλαβαίνει και κατηγοριοποιεί τα δεδομένα δημιουργώντας έτσι απολύτως δομημένα στοιχεία και κατ’ επέκταση ποιο γρήγορη και εύκολη περιήγηση και αναζήτηση.

SPARQL : Semantic Protocol and RDF Query Language, είναι η γλώσσα προγραμματισμού που εκτελεί ερωτήματα στις δομές RDF και επιστρέφει αποτελέσματα. Ακολουθεί την ίδια λογική με την γλώσσα βάσεων δεδομένων SQL, αλλά για διαδικτυακές διασυνδεδεμένες βάσεις δεδομένων.

Topics: Είναι τα θέματα τα οποία αποτελούνται από τρεις ιδιότητες. Subject, Predicate, Object.

Topic map: Είναι ένα πρότυπο αναπαράστασης της σύνδεσης των δεδομένων. Αποτελούνται από τα topics που είναι τα αρχειοθετημένα στοιχεία και έγγραφα, τα associations που είναι οι ενώσεις και οι σχέσεις μεταξύ των topics και τα occurrences που είναι οι πληροφορίες και πηγές που συνδέονται με το κάθε topic. Έτσι δημιουργείται ένας χάρτης εννοιολογικός που εξυπηρετεί άμεσα την τεχνολογία του σημασιολογικού ιστού.

Wandora: Είναι μια εφαρμογή (πλατφόρμα) εξαγωγής και διαχείρισης πληροφοριών από πολλές μορφές διασυνδεδεμένων βάσεων δεδομένων όπως XML/RDF ή OBO αλλά και από διαδικτυακές εφαρμογές που κάνουν χρήση των ανοιχτών διασυνδεδεμένων δεδομένων και αποδίδει γραφικά τα topic maps καθώς και πολλές ακόμα χρήσιμες δομές δεδομένων.