|  |  |
| --- | --- |
| Εικόνα που περιέχει εμπρόσθια όψη, κέρμα  Περιγραφή που δημιουργήθηκε αυτόματα | Εθνικό Μετσόβιο Πολυτεχνείο  Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών  Τομεας Τεχνολογιασ Πληροφορικησ και Υπολογιστων |

**Οι Συζητήσεις Του Κοινοβουλίου Ως Διασυνδεδεμένα Ανοιχτά Δεδομένα**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

**ΠΑΠΑΝΙΚΟΛΑΟΥ ΙΩΑΝΝΗ**

**Επιβλέπων :** Παναγιώτης Τσανάκας

Καθηγητής Ε.Μ.Π.

**Συνεπιβλέπων :** Μάριος Κόνιαρης

Ε.ΔΙ.Π. ΕΜΠ

Αθήνα, Σεπτέμβριος 2023

Η σελίδα αυτή είναι σκόπιμα λευκή.

|  |  |
| --- | --- |
| Εικόνα που περιέχει εμπρόσθια όψη, κέρμα  Περιγραφή που δημιουργήθηκε αυτόματα | Εθνικό Μετσόβιο Πολυτεχνείο  Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών  Τομεας Τεχνολογιασ Πληροφορικησ και Υπολογιστων |

**Οι Συζητήσεις Του Κοινοβουλίου Ως Διασυνδεδεμένα Ανοιχτά Δεδομένα**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

**ΙΩΑΝΝΗ ΠΑΠΑΝΙΚΟΛΑΟΥ**

**Επιβλέπων :** Παναγιώτης Τσανάκας

Καθηγητής Ε.Μ.Π.

**Συνεπιβλέπων :** Μάριος Κόνιαρης

Ε.ΔΙ.Π. ΕΜΠ

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 1η Σεπτεμβρίου 2023.

*(Υπογραφή) (Υπογραφή) (Υπογραφή)*

................................... ................................... ...................................

Παναγιώτης Τσανάκας Βασιλική Καντερέ Ευγενία Τζαννίνη

Καθηγητής Ε.Μ.Π. Επίκουρη Καθηγήτρια Ε.Μ.Π. Καθηγήτρια Ε.Μ.Π.

Αθήνα, Σεπτέμβριος 2023

*(Υπογραφή)*

...................................

**ΙΩΑΝΝΗΣ ΠΑΠΑΝΙΚΟΛΑΟΥ**

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

Copyright © ΙΩΑΝΝΗΣ, ΠΑΠΑΝΙΚΟΛΑΟΥ, 2023.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

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

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

**Περίληψη**

Η αυξανόμενη διαθεσιμότητα ψηφιακών δεδομένων και οι εξελίξεις στις τεχνολογίες σημασιολογικού ιστού έχουν ανοίξει νέες δυνατότητες για την ανάλυση και την εξαγωγή γνώσης από διάφορες πηγές. Η παρούσα διπλωματική εργασία διερευνά την εφαρμογή αυτών των τεχνολογιών για τη μετατροπή των συζητήσεων του Ελληνικού κοινοβουλίου σε συνδεδεμένα δεδομένα, επιτρέποντας πλουσιότερη και πιο ουσιαστική ανάλυση του νομοθετικού λόγου.

Η εργασία ξεκινά με την παρουσία ενός εκτενούς θεωρητικού υποβάθρου. Έπειτα εστιαζόμενο σε πιο πρακτικά ζητήματα, αναλύεται ένα σύστημα το οποίο ξενικά με την μετατροπή των ακατέργαστων κειμενικών δεδομένων των συζητήσεων του Ελληνικού Κοινοβουλίου σε δομημένα έγγραφα XML, τα οποία χρησιμεύουν ως ενδιάμεση αναπαράσταση, επιτρέποντας την περαιτέρω επεξεργασία και ανάλυση. Στη συνέχεια, τα έγγραφα XML μετατρέπονται σε μορφή σχήματος Resource Description Framework Schema (RDFS), επιτρέποντας την αναπαράσταση εννοιών, σχέσεων και μεταδεδομένων που σχετίζονται με τις συζητήσεις.

Εν συνεχεία, στόχος είναι η ανάδειξη της δύναμης των σημασιολογικών ερωτημάτων με τη χρήση της SPARQL, μιας γλώσσας ερωτημάτων για την αναζήτηση δεδομένων RDF. Με τη μετατροπή των συζητήσεων του Ελληνικού Κοινοβουλίου σε διασυνδεδεμένα δεδομένα και την παροχή ενός σχήματος RDFS, καθίσταται δυνατή η εκτέλεση σύνθετων ερωτημάτων, η ανάκτηση ουσιαστικών πληροφοριών και η δημιουργία συνδέσεων μεταξύ διαφόρων πτυχών των συζητήσεων. Η προσέγγιση αυτή επιδεικνύει τις δυνατότητες των τεχνολογιών του σημασιολογικού ιστού στην ενίσχυση της προσβασιμότητας, της ευρεσιμότητας και της χρηστικότητας των κοινοβουλευτικών διαδικασιών, συμβάλλοντας στον ευρύτερο τομέα της ηλεκτρονικής δημοκρατίας και της διαφανούς διακυβέρνησης.

**Λέξεις Κλειδιά:** « Σημασιολογικός Ιστός, Συνδεδεμένα Δεδομένα, XML, RDF, SPARQL, Οι Συζητήσεις Του Ελληνικού Κοινοβουλίου »

Η σελίδα αυτή είναι σκόπιμα λευκή.

**Abstract**

The increasing availability of digital data and developments in semantic web technologies have opened up new possibilities for analysing and extracting knowledge from various sources. This thesis explores the application of these technologies to transform the debates of the Greek parliament into linked data, allowing for richer and more meaningful analysis of legislative discourse.

The paper begins with the presentation of an extensive theoretical background. Then, focusing on more practical issues, a system is analysed which is capable of converting the raw textual data of the Greek Parliament debates into structured XML documents, which serve as an intermediate representation, allowing further processing and analysis. The XML documents are then converted to a Resource Description Framework Schema (RDFS) format, allowing the representation of concepts, relationships and metadata associated with the discussions.

Next, the aim is to demonstrate the power of semantic queries using SPARQL, a query language for querying RDF data. By converting the debates of the Greek Parliament into interlinked data and providing an RDFS schema, it is possible to perform complex queries, retrieve meaningful information and create links between different aspects of the debates. This approach demonstrates the potential of semantic web technologies in enhancing the accessibility, findability and usability of parliamentary procedures, contributing to the broader field of e-democracy and transparent governance.

**Keywords: «**Semantic Web, Linked Data, XML, RDF, SPARQL, Greek Parliament Debates»

Η σελίδα αυτή είναι σκόπιμα λευκή.

Πίνακας περιεχομένων

1 Εισαγωγή 1

1.1 Γενικότερη Ανάλυση Προβλήματος της Εργασίας 1

1.2 Αντικείμενο διπλωματικής 1

1.2.1 Συνεισφορά 2

1.3 Σχετικά Παραδείγματα του Εξωτερικού 1

1.4 Οργάνωση εργασίας 2

2 Θεωρητικό υπόβαθρο 4

2.1 Σημασιολογικός Ιστός - Semantic Web 4

2.2 Ανοιχτά Διασυνδεδεμένα Δεδομένα 4

2.3 Τεχνολογίες Υποστήριξης Διασυνδεδεμένων Δεδομένων 4

2.3.1 XML 4

2.3.2 RDF 4

2.3.3 Λεξιλόγια Σημασιολογικού Ιστού 4

2.3.4 SPARQL 4

2.3 LegalDocML - AkomaNtoso 4

3 Σύστημα Διαχείρισης Πρακτικών 5

3.1 Αρχιτεκτονική Συστήματος 5

3.2 Μέθοδος Αντλήσης Αρχείων - Βάση Δεδομένων 5

3.3 Δομή Πρακτικών Βουλής 5

3.4 Εξαγωγή και Ανάλυση Κειμένων 5

3.4.1 ANLTR4 - REGEΧ 5

3.5 Διαχείριση Δεδομένων 5

3.5.1 Μετατροπή Αρχείων Κειμένου σε XML αρχεία 5

3.5.2 Μετατροπή Αρχείων XML σε RDF αρχεία 5

3.5.3 Επεξεργασία RDF - SPARQL 5

4 Έλεγχος Συστήματος 11

4.1 Μεθοδολογία ελέγχου 11

4.2 Αναλυτική παρουσίαση ελέγχου/παραδείγματος 11

5 Επίλογος 12

5.1 Σύνοψη και συμπεράσματα 12

5.2 Μελλοντικές επεκτάσεις 12

7 Βιβλιογραφία 13

# Εισαγωγή

## Περιγραφή Προβλήματος

~~Εδώ αυτή κάνουμε μια γενική περιγραφή του χώρου εφαρμογής της διπλωματικής. Αναφέρουμε τα χαρακτηριστικά του χώρου και καταλήγουμε στα γενικότερα προβλήματα που αντιμετωπίζει ο χώρος. Η συζήτηση των προβλημάτων θα πρέπει να προϊδεάζει τον αναγνώστη για το τι θα προσπαθήσει να αντιμετωπίσει η διπλωματική, χωρίς ακόμα να αναφερόμαστε συγκεκριμένα στο αντικείμενο της διπλωματικής~~.

Η παρούσα διπλωματική εργασία επικεντρώνεται στην αξιοποίηση των αρχών των συνδεδεμένων δεδομένων και των τεχνολογιών του σημασιολογικού ιστού για τη βελτίωση της προσβασιμότητας και της ανάλυσης των συζητήσεων του Ελληνικού Κοινοβουλίου. Το Ελληνικό Κοινοβούλιο, ως το ανώτατο νομοθετικό όργανο στην Ελλάδα, παράγει έναν τεράστιο όγκο πολύτιμων κειμενικών δεδομένων κατά τη διάρκεια των συζητήσεών του. Ωστόσο, οι υπάρχουσες μέθοδοι πρόσβασης και αξιοποίησης αυτών των δεδομένων παρουσιάζουν αρκετές προκλήσεις και περιορισμούς.

Ο ιστότοπος του Ελληνικού Κοινοβουλίου χρησιμεύει ως η κύρια πηγή πληροφοριών σχετικά με τις κοινοβουλευτικές διαδικασίες. Παρέχει πρόσβαση σε απομαγνητοφωνημένα κείμενα και αρχεία των συζητήσεων, επιτρέποντας στους πολίτες, τους ερευνητές και τους πρωταγωνιστές της πολιτικής σκηνής να ενημερώνονται για τις νομοθετικές δραστηριότητες. Ωστόσο, η τρέχουσα δομή του ιστοτόπου υποστηρίζει κυρίως την παραδοσιακή περιήγηση και αναζήτηση με βάση κυρίως την ημερομηνία, χωρίς προηγμένα χαρακτηριστικά για την εξερεύνηση δεδομένων, τη διασύνδεση και τη σημασιολογική αναζήτηση.

Ένα από τα κύρια προβλήματα που αντιμετωπίζει ο υφιστάμενος δικτυακός τόπος είναι η έλλειψη δομημένης αναπαράστασης δεδομένων. Οι συζητήσεις παρουσιάζονται κυρίως ως αδόμητο κείμενο, σε μορφή pdf, word ή txt, γεγονός που καθιστά δύσκολη την εξαγωγή συγκεκριμένων πληροφοριών ή την εκτέλεση ουσιαστικής ανάλυσης σε λεπτομερές επίπεδο. Αυτός ο περιορισμός εμποδίζει την ολοκληρωμένη έρευνα, τον εντοπισμό τάσεων και την ικανότητα να γίνουν συνδέσεις μεταξύ διαφορετικών συζητήσεων, θεμάτων ή συμμετεχόντων/ομιλητών.

Η διπλωματική αυτή αποσκοπεί στην αντιμετώπιση αυτών των προκλήσεων προτείνοντας ένα πλαίσιο που μετατρέπει τις συζητήσεις του Ελληνικού Κοινοβουλίου σε συνδεδεμένα δεδομένα χρησιμοποιώντας τεχνολογίες σημασιολογικού ιστού. Με την αξιοποίηση της γλώσσας προγραμματισμού Python και του ANTLR-4, τα μη δομημένα κειμένα θα αναλυθούν και θα μετατραπούν σε δομημένα έγγραφα XML. Στη συνέχεια, τα έγγραφα XML θα μετατραπούν σε μορφή σχήματος RDFS (Resource Description Framework Schema), επιτρέποντας την αναπαράσταση εννοιών, σχέσεων και μεταδεδομένων που σχετίζονται με τις συζητήσεις.

Ο πρωταρχικός στόχος είναι να δημιουργηθεί μια αναπαράσταση συνδεδεμένων δεδομένων των συζητήσεων του Ελληνικού Κοινοβουλίου, η οποία θα επιτρέπει τη σημασιολογική αναζήτηση μέσω SPARQL. Αυτή η προσέγγιση θα διευκολύνει τις προηγμένες αναζητήσεις, τα σύνθετα ερωτήματα και την εξαγωγή ουσιαστικών πληροφοριών από το σύνολο δεδομένων. Επιπλέον, η εργασία θα διερευνήσει τη δημιουργία ενός σχήματος RDFS ειδικά για τον τομέα των κοινοβουλευτικών συζητήσεων της Ελλάδος, αποτυπώνοντας την απαραίτητη γνώση για ολοκληρωμένη ανάλυση.

Συνολικά, η παρούσα διπλωματική επιδιώκει να συμβάλει στον τομέα της ηλεκτρονικής δημοκρατίας και της διαφανούς διακυβέρνησης αναδεικνύοντας τις δυνατότητες των τεχνολογιών του σημασιολογικού ιστού στη μετατροπή της προσβασιμότητας, της ανάλυσης και της αξιοποίησης των κοινοβουλευτικών συζητήσεων.

## Αντικείμενο διπλωματικής

Εδώ αναφερόμαστε συγκεκριμένα στο τί θα κάνει η διπλωματική. Αναφέρουμε λεπτομερώς α) τα προβλήματα που θα λύσει (και που ήδη έχουν περιγραφεί γενικά στην προηγούμενη ενότητα), και β) πώς σκοπεύει να τα λύσει.

Είναι σημαντικό κάποιος που θα διαβάσει την ενότητα αυτή να καταλάβει σε σημαντικό βαθμό τον σκοπό της διπλωματικής σας και τις τεχνικές δυσκολίες της, χωρίς να είναι αναγκαίο να δει όλα τα άλλα κεφάλαια. **Η ενότητα αυτή θέλει πολύ προσοχή και καλύτερα να τη γράψετε αφού έχετε γράψει όλα τα υπόλοιπα κεφάλαια.**

### Συνεισφορά

Εδώ παραθέτουμε αριθμητικά συγκεκριμένες ενέργειες/λύσεις/μεθοδολογίες που παρουσιάζει η διπλωματική και λύνουν τα προβλήματα που υποσχεθήκαμε στην προηγούμενη ενότητα ότι θα λύσει η διπλωματική. Συνήθως η υποενότητα αυτή έχει την παρακάτω μορφή:

Η συνεισφορά της διπλωματικής συνοψίζεται ως εξής:

1. Μελετήσαμε συστήματα κ.λ.π.
2. Υλοποιήσαμε 3 αλγορίθμους υπολογισμού κ.λ.π.
3. Αξιολογήσαμε την επίδοση των αλγορίθμων και βρήκαμε ότι κ.λ.π.
4. Ενσωματώσαμε τους αλγορίθμους σε πρότυπο σύστημα κ.λ.π.
5. ...

## Σχετικά Παραδείγματα του Εξωτερικού

Πρακτικά Ευρωπαϊκού Κοινοβουλίου <https://linkedpolitics.ops.few.vu.nl/web/html/home.html>

Πρακτικά Κοινοβουλίου του Καναδά

<https://hansard.opennwt.ca/debates/>

## Οργάνωση κειμένου

Εδώ περιγράφουμε τα κεφάλαια της διπλωματικής: 1 πρόταση για το τι θα έχει κάθε κεφάλαιο. Συνήθως η ενότητα αυτή έχει την παρακάτω μορφή (δεν θα σας πάρει πάνω από 1 μεγάλη παράγραφο):

Εργασίες σχετικές με το αντικείμενο της διπλωματικής παρουσιάζονται στο Κεφάλαιο 2 . Το Κεφάλαιο 3 συζητά θέματα μοντελοποίησης. Στο Κεφάλαιο 4 αναπτύσσουμε κ.λ.π.

# Θεωρητικό υπόβαθρο

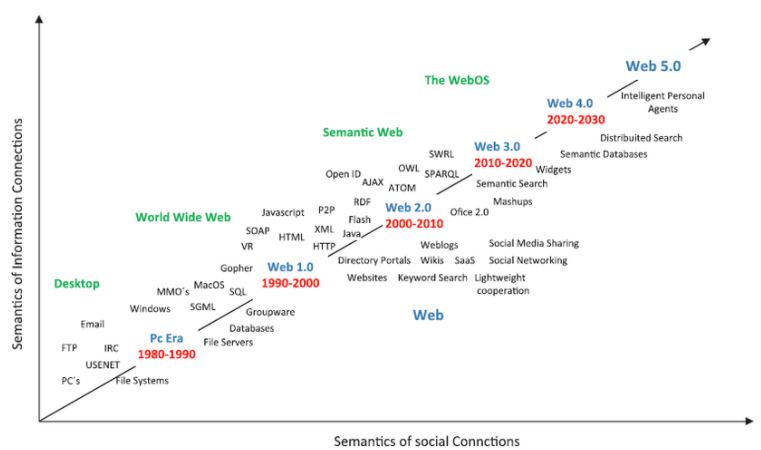
~~Εδώ γράφουμε σύντομα τις τεχνικές/μεθοδολογίες/μοντέλα που πιθανά θα χρησιμοποιήσει η διπλωματική και είναι αναγκαία η κατανόησή τους από τον αναγνώστη πριν από την παρουσίαση της ανάλυσης και σχεδίασης του συστήματος.~~

~~Πρόκειται για τεχνικές/μεθοδολογίες/μοντέλα που έχουν προταθεί από άλλους και δεν είναι πρωτότυπη δουλειά της διπλωματικής.~~

~~Μετά βάζουμε 1 ενότητα για κάθε τεχνική/μεθοδολογία/μοντέλο, όπου και δίνουμε λεπτομερή περιγραφή.~~

## Σημασιολογικός Ιστός – Semantic Web

Με την πάροδο των χρόνων, η ανάπτυξη του διαδικτύου έχει φέρει επανάσταση στον τρόπο με τον οποίο επικοινωνούμε, έχουμε πρόσβαση σε πληροφορίες και διεξάγουμε επιχειρηματικές δραστηριότητες. Από τα πρώτα βήματα του, ως ένα απλό δίκτυο διασυνδεδεμένων εγγράφων, το διαδίκτυο έχει υποστεί αξιοσημείωτες αλλαγές με την πάροδο των ετών. Μια τέτοια «τελευταία» αλλαγή είναι η μετάβαση στον σημασιολογικό ιστό, γεγονός που αποσκοπεί στη βαθύτερη κατανόηση και νοηματοδότηση του περιεχομένου του ιστού. Αυτή η μετατόπιση έχει ανοίξει νέες δυνατότητες για την αναπαράσταση γνώσης, την ολοκλήρωση δεδομένων και την ευφυή αυτοματοποίηση.

Σχήμα 2-1: Η εξέλιξη του Διαδικτύου (Osorio, Ortiz 2013 ‐ <https://myeltcafe.com/wp-content/uploads/2020/12/123123123.jpg> )

Η μετάβαση από τον παραδοσιακό ιστό στον σημασιολογικό ιστό αποτελεί σημαντικό ορόσημο στην εξέλιξη του διαδικτύου. Στα πρώτα στάδια του ιστού, οι πληροφορίες παρουσιάζονταν κυρίως σε αδόμητες μορφές, γεγονός που καθιστούσε δύσκολη έως και αδύνατη την αποτελεσματική επεξεργασία και ερμηνεία του περιεχομένου από τις μηχανές. Ωστόσο, καθώς ο ιστός μεγάλωνε σε μέγεθος και πολυπλοκότητα, προέκυψε η ανάγκη για έναν πιο έξυπνο και αποτελεσματικό τρόπο οργάνωσης και κατανόησης του τεράστιου όγκου των διαθέσιμων πληροφοριών. Έτσι γεννιέται η έννοια του σημασιολογικού ιστού, η οποία αποσκοπούσε στο να προσδώσει στο περιεχόμενο του ιστού σαφές νόημα και να επιτρέψει στις μηχανές να κατανοήσουν τα δεδομένα.

Στον πυρήνα του σημασιολογικού ιστού βρίσκεται η χρήση οντολογιών, οι οποίες παρέχουν μια τυπική αναπαράσταση της γνώσης σε συγκεκριμένους τομείς. Οι οντολογίες ορίζουν τις έννοιες, τις σχέσεις και τις ιδιότητες ενός συγκεκριμένου τομέα, διευκολύνοντας την κοινή κατανόηση μεταξύ ανθρώπων και μηχανών. Με την αξιοποίηση των οντολογιών, ο σημασιολογικός ιστός επιτρέπει τον σχολιασμό και τη δόμηση του περιεχομένου του ιστού, καθιστώντας το αναγνώσιμο από μηχανές.

Τα συνδεδεμένα δεδομένα είναι μια άλλη βασική πτυχή του σημασιολογικού ιστού. Περιλαμβάνει τη δημιουργία συνδέσεων μεταξύ διαφορετικών πηγών δεδομένων στον ιστό, χρησιμοποιώντας τυποποιημένα πρωτόκολλα και τεχνολογίες όπως το RDF (Resource Description Framework) και τα URIs (Uniform Resource Identifiers). Συνδέοντας δεδομένα σε διαφορετικούς τομείς και πηγές, ο σημασιολογικός ιστός προωθεί την ολοκλήρωση, τη διασύνδεση και την επαναχρησιμοποίηση δεδομένων. Αυτός ο διασυνδεδεμένος ιστός δεδομένων επιτρέπει μια πιο ολοκληρωμένη και πλαισιωμένη κατανόηση των πληροφοριών, ξεπερνώντας τα όρια των μεμονωμένων εγγράφων ή ιστοτόπων. Φυσικά, οι παραπάνω έννοιες θα αναλυθούν εκτενώς στην συνέχεια.

Προφανώς, η μετάβαση στον σημασιολογικό ιστό υπόσχεται πολλά για διάφορους τομείς. Ενδεικτικά, στον τομέα της ανάκτησης πληροφοριών και των μηχανών αναζήτησης, ο σημασιολογικός ιστός μπορεί να βελτιώσει την ακρίβεια και τη συνάφεια των αποτελεσμάτων αναζήτησης, κατανοώντας το νόημα και το πλαίσιο πίσω από τα ερωτήματα των χρηστών. Σε έναν άλλο τομέα, όπως αυτού του ηλεκτρονικό εμπόριο, ο σημασιολογικός ιστός επιτρέπει πιο έξυπνες συστάσεις προϊόντων και εξατομικευμένες εμπειρίες αγορών με βάση τη βαθύτερη κατανόηση των προτιμήσεων και των αναγκών των πελατών. Τέλος, στην υγειονομική περίθαλψη, ο σημασιολογικός ιστός διευκολύνει τη διαλειτουργικότητα και την ενοποίηση των ιατρικών δεδομένων, οδηγώντας σε βελτιωμένη περίθαλψη των ασθενών, ερευνητική συνεργασία και λήψη αποφάσεων βάσει δεδομένων.

Επιπλέον, ο σημασιολογικός ιστός έχει επιπτώσεις στις εφαρμογές τεχνητής νοημοσύνης (AI) και μηχανικής μάθησης (ML). Παρέχοντας ένα τυποποιημένο πλαίσιο για την αναπαράσταση δεδομένων και την ολοκλήρωση της γνώσης, ο σημασιολογικός ιστός μπορεί να βελτιώσει την εκπαίδευση και την απόδοση των μοντέλων τεχνητής νοημοσύνης, επιτρέποντάς τους να κάνουν πιο τεκμηριωμένες προβλέψεις και συστάσεις. Ανοίγει επίσης το δρόμο για την ανάπτυξη ευφυών πρακτόρων και chatbots που μπορούν να κατανοούν και να απαντούν σε ερωτήματα χρηστών με πιο φυσικό και συνειδητό τρόπο.  
 Πλέον, μοναδικός στόχος της τεχνολογίας είναι η πρόοδος της, επομένως όπως μας μαρτυρά και το Σχήμα 2-1, είναι ότι πλησιάζει η εδραίωση του Semantic Web ή του Web 3.0 και η σταδιακή μας μετάβαση στο νεότερο διαδίκτυο με αρίθμηση 4.0, αν βέβαια επιβεβαιωθούν αυτές οι προβλέψεις, μέχρι το τέλος αυτής της δεκαετίας.

## Ανοιχτά Διασυνδεδεμένα Δεδομένα

Η ιδέα των Ανοικτών Συνδεδεμένων Δεδομένων είναι ένα θεμελιώδες στοιχείο που, εκτός από τον Σημασιολογικό Ιστό, βελτιώνει την αποτελεσματικότητα και τη διαλειτουργικότητα των δεδομένων. Ενθαρρύνοντας την τυποποιημένη και ανοικτή δημοσίευση και σύνδεση δομημένων δεδομένων, τα Ανοικτά Συνδεδεμένα Δεδομένα προωθούν τα ιδανικά του Σημασιολογικού Ιστού.

Προκειμένου να εκφραστούν τα δεδομένα με τρόπο που να είναι αναγνώσιμο από μηχανήματα, τα Ανοιχτά Συνδεδεμένα Δεδομένα επιτάσσουν τη χρήση ανοιχτών προτύπων όπως το RDF (Resource Description Framework). Το RDF επιτρέπει την τριπλή δομή υποκειμένου-κατηγορουμένου-αντικειμένου των δεδομένων, επιτρέποντας τη δήλωση σύνθετων σημασιολογιών και συνδέσεων εντός των δεδομένων. Αυτό διευκολύνει τον συνδυασμό και την ενσωμάτωση συνόλων δεδομένων από πολλές πηγές, δημιουργώντας ένα δίκτυο γνώσης που είναι πιο εκτεταμένο και συνδεδεμένο.

Η χρήση ομοιόμορφων αναγνωριστικών πόρων (URI) για τον ειδικό προσδιορισμό πραγμάτων και εννοιών εντός του οικο-συστήματος των συνδεδεμένων δεδομένων επιβάλλεται επίσης από τις αρχές των ανοικτών συνδεδεμένων δεδομένων. Τα URI χρησιμεύουν ως μόνιμα και παγκοσμίως μοναδικά αναγνωριστικά, επιτρέποντας την απρόσκοπτη αναφορά και σύνδεση πόρων σε διαφορετικά σύνολα δεδομένων και τομείς.

Όλα όσα περιγράψαμε παραπάνω έχει συμπυκνώσει ο Τιμ Μπέρνερς Λι, ο εφευρέτης του Παγκόσμιου Ιστού, σε τέσσερις βασικές και θεμελιώδης αρχές. Πιο συγκεκριμένα:

- Για να διασφαλιστεί ότι κάθε οντότητα ή έννοια έχει μια διακριτή ταυτότητα που την ξεχωρίζει μέσα στο τεράστιο διασυνδεδεμένο δίκτυο δεδομένων, ο πρώτος κανόνας απαιτεί τη χρήση URIs ως ονόματα για τα πράγματα.

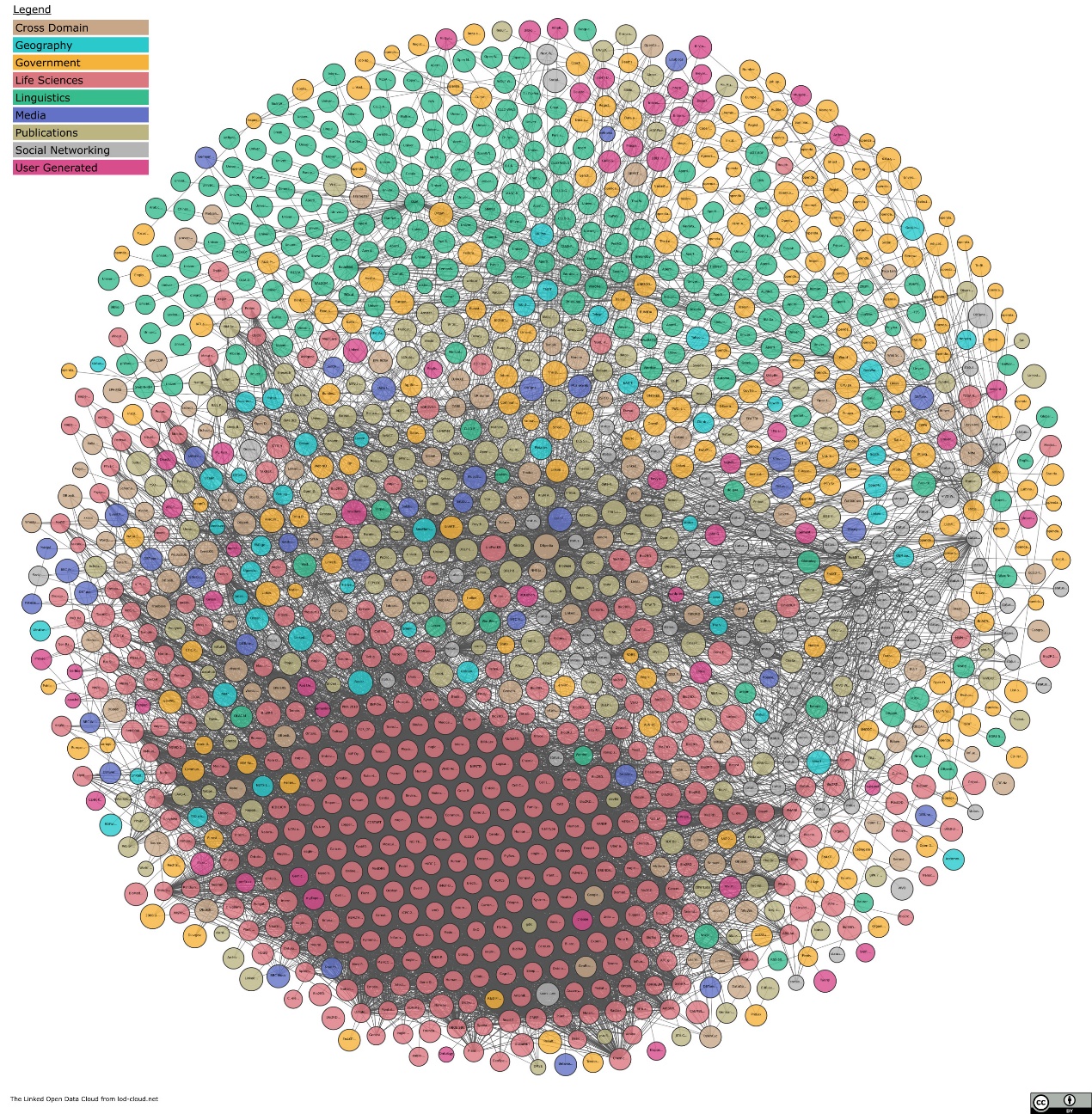
- Ο δεύτερος κανόνας καθιστά απλή την αναζήτηση αυτών των ονομάτων με τη χρήση HTTP URIs, επιτρέποντας στους χρήστες να έχουν πρόσβαση και να ανακτούν χρήσιμα δεδομένα που σχετίζονται με τους αναγνωρισμένους πόρους.

- Για την επιτυχή διευθέτηση και αναζήτηση των δεδομένων, ο τρίτος κανόνας δίνει έμφαση στην παρουσίαση σχετικών πληροφοριών κατά την αναζήτηση ενός URI. Για το σκοπό αυτό χρησιμοποιούνται τυποποιημένες τεχνολογίες όπως το RDF και η SPARQL.

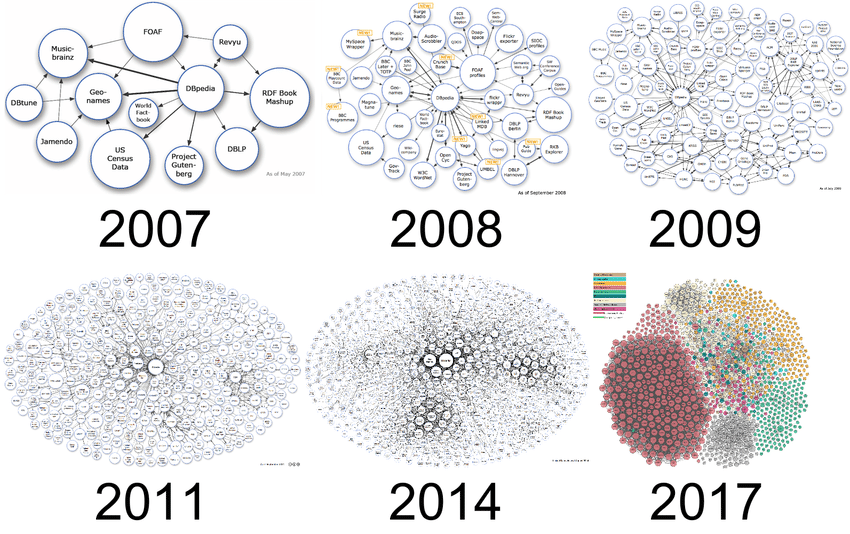
- Η συμπερίληψη συνδέσμων προς άλλα URI τονίζεται επίσης από το τέταρτο κριτήριο, διευκολύνοντας την εύρεση νέων σχετικών πόρων και την ανάπτυξη ενός δικτύου διασυνδεδεμένης γνώσης.

Η τήρηση αυτών των κατευθυντήριων γραμμών επιτρέπει απροσδόκητες επαναχρησιμοποιήσεις και καινοτόμες ανακαλύψεις που ωθούν την καινοτομία και προάγουν τη μεγαλύτερη κατανόηση του κόσμου, απελευθερώνοντας το πλήρες δυναμικό των συνδεδεμένων δεδομένων.

Τα συνδεδεμένα δεδομένα λοιπόν είναι μια έννοια που προωθεί τη διασύνδεση και την ενσωμάτωση διαφορετικών συνόλων δεδομένων στο διαδίκτυο, όπως φαίνεται από το Σχήμα 2-1 (LOD Cloud). Στο διάγραμμα αυτό αποτυπώνεται το μεγαλειώδες δίκτυο των συνδεδεμένων συνόλων δεδομένων, τα οποία εμφανίζονται ως κόμβοι, και τις συνδέσεις τους. Ένα σύνολο δεδομένων σύμφωνα με τις αρχές που είδαμε παραπάνω, αναπαρίσταται από κάθε κόμβο στο LOD Cloud, το οποίο χρησιμοποιεί μοναδικά αναγνωριστικά (URIs) για τον προσδιορισμό και την αναφορά των πόρων. Οι σύνδεσμοι και οι συνδέσεις μεταξύ των διαφόρων συνόλων δεδομένων αναπαρίστανται από τους συνδέσμους μεταξύ των κόμβων, επιτρέποντας την εύκολη πλοήγηση και την εξερεύνηση πληροφοριών. Το διάγραμμα της εικόνας δείχνει τη δύναμη της διασύνδεσης δεδομένων μεταξύ τους, επιτρέποντας στους χρήστες να περιηγηθούν μεταξύ συνόλων δεδομένων, να βρουν νέα δεδομένα και να αποκτήσουν ολοκληρωμένες γνώσεις, αξιοποιώντας τη συλλογική σοφία που υπάρχει στο οικοσύστημα των συνδεδεμένων δεδομένων.

Σχήμα 2-1 The Linked Open Data Cloud (<http://cas.lod-cloud.net/clouds/lod-cloud-sm.jpg>)

Ακόμα ένα άλλο ενδιαφέρον γράφημα είναι αυτό του Σχήματος 2-2 στο οποίο απεικονίζεται η εκθετική ανάπτυξη των Συνδεδεμένων Ανοικτών Δεδομένων (LOD) από το 2007 έως και τα τελευταία χρόνια, αναδεικνύοντας την αξιοσημείωτη πρόοδο που έχουν σημειώσει τα Διασυνδεδεμένα Δεδομένα με την πάροδο του χρόνου. Το γράφημα καταγράφει το αυξανόμενο δίκτυο πόρων συνδεδεμένων δεδομένων, δείχνοντας την αύξηση του αριθμού των συνόλων δεδομένων που δημοσιεύονται ως LOD. Η επέκταση αυτή αντανακλά την αυξανόμενη αποδοχή και εκτίμηση των αρχών των συνδεδεμένων δεδομένων σε ένα ευρύ φάσμα πεδίων και τομέων. Η ποσότητα και η ποικιλία των διασυνδεδεμένων συνόλων δεδομένων έχουν αυξηθεί τρομερά καθώς όλο και περισσότερες επιχειρήσεις, κοινωνίες και κυβερνήσεις υιοθετούν την ιδέα των συνδεδεμένων δεδομένων. Η ανάπτυξη ενός τεράστιου, διασυνδεδεμένου γράφου γνώσης ως αποτέλεσμα αυτής της ανάπτυξης κατέστησε δυνατή την καλύτερη ενσωμάτωση, ανακάλυψη και χρήση δεδομένων.



Σχήμα 2-2 Ανάπτυξη των συνδεδεμένων ανοικτών δεδομένων από το 2007 (<https://www.researchgate.net/figure/Growth-of-Linked-Open-Data-since-2007-1-The-amount-of-data-sets-published-as-LOD-have_fig2_331748480> )

## Τεχνολογίες Υποστήριξης Δεδομένων

Στον τομέα του Σημασιολογικού Ιστού και των Συνδεδεμένων Δεδομένων περιλαμβάνονται διάφορες τεχνολογίες που είναι ζωτικής σημασίας για την αποτελεσματική αναπαράσταση, ολοκλήρωση, επερώτηση και διαλειτουργικότητα των δεδομένων. Οι τεχνολογίες αυτές προσφέρουν το πλαίσιο για την οργάνωση, τη σύνδεση και τη συλλογή γνώσης από διάφορα σύνολα δεδομένων. Μια ευέλικτη και ευρέως χρησιμοποιούμενη γλώσσα για την κωδικοποίηση δομημένων δεδομένων, η XML (eXtensible Markup Language) προωθεί την ανταλλαγή δεδομένων και τη διαλειτουργικότητα. Προκειμένου να καταστεί δυνατή η ανάπτυξη πλούσιων σημασιολογικών αναπαραστάσεων, το RDF (Resource Description Framework) παρουσιάζει ένα τυποποιημένο μοντέλο για την περιγραφή και τη σύνδεση δεδομένων. Ο πυρήνας των συνδεδεμένων δεδομένων, το RDF προσφέρει έναν αποτελεσματικό τρόπο περιγραφής σχέσεων, ιδιοτήτων και πληροφοριών. Με τη βοήθεια της ισχυρής γλώσσας ερωτημάτων SPARQL (SPARQL Protocol and RDF Query Language), οι χρήστες μπορούν να λάβουν συγκεκριμένα δεδομένα, να πραγματοποιήσουν περίπλοκες συνδέσεις και να αποκτήσουν κατανόηση από τα διασυνδεδεμένα δεδομέν. Αυτές οι τεχνολογίες συνεργάζονται για να δημιουργήσουν μια πλήρη εργαλειοθήκη που επιτρέπει στους ερευνητές, τους προγραμματιστές και τους επαγγελματίες των δεδομένων να αξιοποιούν πλήρως τα έργα του σημασιολογικού ιστού και των συνδεδεμένων δεδομένων.

### Extensible Markup Language (XML)

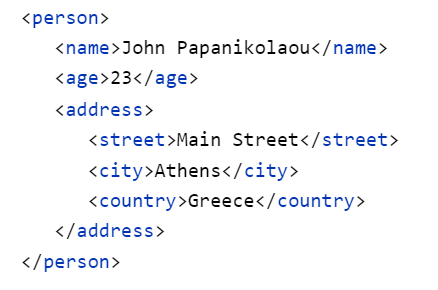
Η eXtensible Markup Language, γνωστή ως XML, είναι μια ευρέως υιοθετημένη τεχνολογία στον τομέα του Σημασιολογικού Ιστού και των συνδεδεμένων δεδομένων, καθώς είναι μια ισχυρή γλώσσα σήμανσης που έχει επηρεάσει σημαντικά την ανάπτυξη της ανταλλαγής πληροφοριών στο διαδίκτυο. Η αφετηρία της εντοπίζεται στις αρχές της δεκαετίας του 1970, όταν κατέστη για πρώτη φορά αναγκαίος ένας δομημένος τρόπος αναπαράστασης και ανταλλαγής δεδομένων.

Στις δεκαετίες του 1980 και 1990 δημιουργήθηκαν διάφορες γλώσσες σήμανσης (όπως η HTML), η καθεμία με μοναδικούς περιορισμούς και προβλεπόμενες χρήσεις. Ωστόσο, για να αντιμετωπιστούν οι αυξανόμενες απαιτήσεις για ανταλλαγή δεδομένων σε διάφορες πλατφόρμες και συστήματα, απαιτήθηκε μια πιο προσαρμόσιμη και επεκτάσιμη λύση, με αποτέλεσμα στα τέλη της δεκαετίας του 1990, να αναπτυχθεί η XML.

Η Κοινοπραξία του Παγκόσμιου Ιστού (W3C) ξεκίνησε τη διαδικασία τυποποίησης της XML το 1996 και καθόρισε τη σύνταξη και τη σημασιολογία της. Ο κύριος στόχος ήταν η ανάπτυξη μιας γλώσσας που θα επέτρεπε στους χρήστες να κατασκευάζουν τις δικές τους μοναδικές γλώσσες σήμανσης, καθιστώντας την αρκετά ευέλικτη ώστε να μπορεί να εφαρμοστεί σε ένα ευρύ φάσμα εφαρμογών και δομών δεδομένων.

Η κύρια καινοτομία της XML είναι το πόσο απλή και κατανοητή είναι. Τα στοιχεία περιέχονται σε αγκύλες και διατάσσονται ιεραρχικά χρησιμοποιώντας μια μέθοδο βασισμένη σε ετικέτες. Τόσο οι άνθρωποι όσο και οι μηχανές μπορούν να ερμηνεύσουν τα δεδομένα, δεδομένου ότι αυτές οι ετικέτες καθορίζουν τη δομή και το νόημά τους.

Παρακάτω παρατίθεται ένα απόσπασμα από μια πιθανή απεικόνιση της XML με χρήση ετικετών για να γίνει κατανοητή η σύνταξη της γλώσσας:



Σχήμα 2.3.1 Δείγμα Κώδικα XML

Στο παράδειγμα του σχήματος 2.3.1, οι ετικέτες XML ενσωματώνουν διαφορετικά κομμάτια πληροφοριών για ένα άτομο. Το βασικό στοιχείο (root) είναι η ετικέτα "person", η οποία περιέχει φωλιασμένες ετικέτες όπως "name", "age" και "address". Κάθε ετικέτα αντιπροσωπεύει ένα μοναδικό κομμάτι δεδομένων. Για παράδειγμα, οι ετικέτες "name" και "age" περιέχουν το όνομα και την ηλικία του ατόμου αντίστοιχα, ενώ η ετικέτα "address" έχει φωλιασμένες τις ετικέτες "street", "city" και "country" για να αντικατοπτρίζει τις πληροφορίες διεύθυνσης.

Η ανεξαρτησία πλατφόρμας, η οποία επιτρέπει τη χρήση της XML σε πολλές πλατφόρμες και γλώσσες υπολογιστών, είναι ένα από τα κύρια πλεονεκτήματά της. Εξαιτίας αυτού, η XML είναι η καλύτερη μορφή που μπορεί να χρησιμοποιηθεί για την ανταλλαγή δεδομένων και τη διαλειτουργικότητα. Τα έγγραφα XML μπορούν επίσης να διαβαστούν από ανθρώπους, γεγονός που βελτιώνει την κατανόηση και την προσβασιμότητά τους τόσο από μηχανές όσο και από ανθρώπους.

Στο πλαίσιο των συνδεδεμένων δεδομένων, η XML αποτελεί βασική τεχνική για τη δόμηση και τη δημιουργία εγγράφων αναγνώσιμων από μηχανήματα για μη δομημένα κείμενα. Αυτά τα έγγραφα μπορούν να επεξεργαστούν, να ταξινομηθούν και να επισημανθούν με χρήσιμα στοιχεία και χαρακτηριστικά με τη χρήση XML, με αποτέλεσμα την παραγωγή καλά διαμορφωμένων εγγράφων XML. Τα επόμενα βήματα της μετάφρασης, της ολοκλήρωσης και της αναζήτησης δεδομένων με τη χρήση διαφόρων τεχνολογιών σημασιολογικού ιστού καθίστανται δυνατά με αυτή τη διαδικασία.

Σε καταστάσεις που αφορούν την ενσωμάτωση δεδομένων, όπου διάφορες πηγές πρέπει να εναρμονιστούν σε μια ενιαία αναπαράσταση, η XML είναι επίσης απαραίτητη. Είναι απλούστερη η αντιστοίχιση και ο μετασχηματισμός διαφορετικών συνόλων δεδομένων σε μια ενιαία μορφή λόγω της καθολικής σύνταξης της XML για την αναπαράσταση δομών δεδομένων.

### Resource Description Framework (RDF)

### Λεξιλόγια Σημαντικού Ιστού

Foaf, dcterms κτλ

### SPARQL

## LegalDocMl – Akoma Ntoso

# Ανάλυση Απαιτήσεων Συστήματος

Εδώ λέμε ότι θα ακολουθήσει η περιγραφή της αρχιτεκτονικής του συστήματος και θα γίνει η ανάλυση απαιτήσεων για τις λειτουργίες του.

## Αρχιτεκτονική

Εδώ παρουσιάζουμε τα επιμέρους κομμάτια (υποσυστήματα) από τα οποία θεωρούμε ότι αποτελείται το σύστημά μας. Για κάθε υποσύστημα δίνεται μια σύντομη περιγραφή (η αναλυτική περιγραφή των λειτουργιών του ακολουθεί στην επόμενη ενότητα). Τέλος, δίνουμε ένα γενικό σχήμα που δείχνει τα υποσυστήματα και πώς αυτά επικοινωνούν μεταξύ τους. Το σχήμα αυτό αρκεί να είναι απλό block diagram και όχι Διάγραμμα Ροής Δεδομένων (DFD).

## Περιγραφή Λειτουργιών

Εδώ περιγράφουμε τις λειτουργίες που απαιτείται να εκτελεί το σύστημα. Αφού έχουμε χωρίσει ήδη το σύστημά μας σε υποσυστήματα, περιγράφουμε το κάθε υποσύστημα ξεχωριστά. Κάθε περιγραφή υποσυστήματος έχει 2 σκέλη: κείμενο και σχήμα. α) Για το κείμενο: με απλά λόγια και πραγματικά παραδείγματα όπου χρειάζεται, πρέπει να εξηγήσουμε στο χρήστη τί είναι αυτό που κάνει το υποσύστημα. Καλό είναι οι εξηγήσεις να δίνονται σύντομα, περιεκτικά και αριθμημένα, π.χ. 1. αυτό, 2. το άλλο, κ.λ.π.. β) Για το σχήμα: μετά από το κείμενο, έχουμε και σχετικό Διάγραμμα Ροής Δεδομένων (DFD) με το οποίο δίνουμε και σχηματικά τί κάνει το υποσύστημα.

Είναι πολύ βασικό να κατανοήσετε ότι εδώ περιγράφουμε **τί κάνει** το σύστημα **και ΟΧΙ πώς** το κάνει.Για το λόγο αυτό, να χρησιμοποιείτε και αρκετά παραδείγματα.

### <Τίτλος υποσυστήματος 1>

<<……>>

### <Τίτλος υποσυστήματος 2>

<<……>>

## Μοντέλο Οντοτήτων Συσχετίσεων

Αν για το σύστημα μας σκοπεύουμε να χρησιμοποιήσουμε Βάση Δεδομένων, δίνουμε εδώ το μοντέλο Οντοτήτων- Συσχετίσεων.

# Σχεδίαση Συστήματος

Εδώ λέμε ότι θα ακολουθήσει η σχεδίαση του συστήματος.

## Αρχιτεκτονική

Εδώ παρουσιάζουμε τα επιμέρους κομμάτια από τα οποία θεωρούμε ότι έχει κτιστεί ο κώδικάς μας. Συνήθως, επειδή οι περισσότεροι γράφετε σε αντικειμενοστρεφή γλώσσα προγραμματισμού (C++/JAVA), τα κομμάτια είναι στην ουσία οι κλάσεις της εφαρμογής. Για κάθε κλάση γράψτε μια σύντομη περιγραφή (η αναλυτική περιγραφή των μεθόδων/λειτουργιών/συναρτήσεων του θα ακολουθήσει στην επόμενη ενότητα). Τέλος, δίνουμε ένα γενικό σχήμα που δείχνει τις κλάσεις και πώς αυτές επικοινωνούν μεταξύ τους. Το σχήμα αυτό π.χ. αρκεί να είναι ένα απλό block diagram κλάσεων, όπου θα φαίνεται η κληρονομικότητα και οι συνδέσεις μεταξύ των κλάσεων. Μια κλάση συνδέεται με μια άλλη αν μια μέθοδός της χρησιμοποιεί αντικείμενο από την άλλη ως παράμετρο.

## Περιγραφή Κλάσεων

Εδώ περιγράφουμε συνοπτικά τις λειτουργίες/μεθόδους/συναρτήσεις των κλάσεων. Καλό είναι οι περιγραφές να δίνονται σύντομα, περιεκτικά και αριθμημένα, π.χ. 1. αυτό, 2. το άλλο, κ.λ.π. Δεν βάζουμε κώδικα καθόλου!

### <Τίτλος κλάσης 1>

<<……>>

### <Τίτλος κλάσης 2>

<<……>>

## Βάση Δεδομένων

Αν το σύστημα μας χρησιμοποιεί Βάση Δεδομένων, περιγράφουμε το σχήμα της βάσης που προκύπτει από το μοντέλο Οντοτήτων-Συσχετίσεων που είχαμε στο προηγούμενο κεφάλαιο.

## Κωδικοποίηση αρχείων

Αν το σύστημα χρησιμοποιεί αρχεία στα οποία αποθηκεύεται κωδικοποιημένα κάποια πληροφορία, θα πρέπει να δοθεί εδώ η περιγραφή της κωδικοποίησής τους.

# Υλοποίηση

Εδώ λέμε ότι θα συζητήσουμε λεπτομερώς θέματα υλοποίησης του συστήματος.

## Λεπτομέρειες υλοποίησης

Εδώ περιγράφουμε λεπτομερώς θέματα της διπλωματικής που έχουν τεχνικό ή αλγοριθμικό ενδιαφέρον. Για παράδειγμα, κάποια μέθοδος σε κάποια κλάση ενδεχομένως να υλοποιεί έναν πολύπλοκο αλγόριθμο. Εδώ είναι το κατάλληλο σημείο για την περιγραφή του.

Προσδιορίστε επομένως τα θέματα αυτά, βάλτε μια ενότητα για κάθε ένα και περιγράψτε τα αναλυτικά. Η περιγραφή μπορεί να γίνει βάζοντας κομμάτια κώδικα ή ψευδοκώδικα, και περιγράφοντάς τα με λόγια. Μην ξεχνάτε να δίνετε πάντα παραδείγματα για το πώς τρέχει ένα κομμάτι κώδικα π.χ. για έναν αλγόριθμο.

### <Τίτλος θέματος 1>

<<……>>

### <Τίτλος θέματος 2>

<<……>>

## Πλατφόρμες και προγραμματιστικά εργαλεία

Εδώ περιγράφονται τα χαρακτηριστικά της συγκεκριμένης υλοποίησης, όπως η πλατφόρμα ανάπτυξης και εκτέλεσης, τα προγραμματιστικά εργαλεία, οι απαιτήσεις της εφαρμογής σε hardware, κ.λ.π.

Επίσης, περιγράφεται λεπτομερώς η διαδικασία εγκατάστασης της διπλωματικής σε υπολογιστή. Προσέξτε να δίνονται όλες οι λεπτομέρειες, το απαραίτητο λογισμικό και οι αναγκαίες ρυθμίσεις.

# Έλεγχος

Εδώ λέμε ότι θα ακολουθήσει η αξιολόγηση του συστήματος.

## Μεθοδολογία ελέγχου

Στην ενότητα αυτή λέμε ότι ο έλεγχος πραγματοποιήθηκε με τη χρήση ενός σεναρίου λειτουργίας. Περιγράφουμε σύντομα το σενάριο.

## Αναλυτική παρουσίαση ελέγχου

Στην ενότητα αυτή παρουσιάζουμε αναλυτικά τον έλεγχο του συστήματος σύμφωνα με το σενάριο που περιγράφηκε στην προηγούμενη ενότητα.

Το κεφάλαιο αυτό, λειτουργεί και ως εγχειρίδιο χρήσης του προγράμματος.

# Επίλογος

Εδώ λέμε ότι θα συνοψίσουμε την παρουσίαση της διπλωματικής.

## Σύνοψη και συμπεράσματα

Εδώ συνοψίζουμε τα αποτελέσματα της διπλωματικής και περιγράφουμε τα συμπεράσματα που προέκυψαν, αρνητικά και θετικά. Επιβεβαιώνουμε τη συνεισφορά της διπλωματικής στα προβλήματα που αναφέραμε στην εισαγωγή.

## Μελλοντικές επεκτάσεις

Εδώ δίνουμε ιδέες για επέκταση της διπλωματικής.

# Βιβλιογραφία

|  |  |
| --- | --- |
| [BBC+99] | P.A. Bernstein, Th. Bergstraesser, J. Carlson, S. Pal, P. Sanders, D. Shutt. Microsoft Repository Version 2 and the Open Information Model. To appear in Information Systems 24(2), 1999. |
| [BCR94] | V. R. Basili, G.Caldiera, H. D. Rombach. The Goal Question Metric Approach. Encyclopedia of Software Engineering - 2 Volume Set, pp. 528-532, John Wiley & Sons, Inc., available at http://www.cs.umd.edu/users/basili/papers.html, 1994 |
| [Dea97] | E. B. Dean, "Quality Functional Deployment from the Perspective of Competitive Advantage", available at http://mijuno.larc.nasa.gov/dfc/qfd.html |
| [JJQV98] | M. Jarke, M.A.Jeusfeld, C. Quix, P. Vassiliadis: Architecture and quality in data warehouses, Proceedings CΑiSE 98, Pisa, Italy, 1998. |
| [JV97] | M. Jarke, Y. Vassiliou. Foundations of data warehouse quality – a review of the DWQ project. In Proc. 2nd Intl. Conference Information Quality (IQ-97), Cambridge, Mass., 1997. |
| [Orr98] | K. Orr. Data quality and systems theory. In Communications of the ACM, 41, 2, pp. 54-57, Feb. 1998. |